Dealing with imperfections: from aberration to scattering

Thesis by
Ruizhi Cao

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2024
Defended December 21, 2023

ACKNOWLEDGEMENTS

The journey of working as a Ph.D. student can be both rewarding and demanding. We aim to obtain new knowledge and develop new techniques in our own research, which can become challenging from time to time. This would not have been possible for me without the generous help and support of a number of people.

I would like to first thank my advisor, Prof. Changhuei Yang, who generously offered me the chance to explore some interesting problems even if we had a hardship in finding the funding for those fruitful studies. Many of my works would not have survived without his support. His attitude and standards on research also influence me in identifying the most valuable questions in research.

I would like to thank my committee, Prof. Lihong Wang, Prof. Palghat P. Vaidyanathan, and Prof. Alireza Marandi, for their suggestions and advices on my research. I took several courses from them and I learned a lot from their lectures. Their passion for their own research and lectures also motivated me during my hardships.

I would like to thank Dr. Yan Liu and Dr. Michelle Cua, who taught me various experiment skills I can use in my own studies. Without their help, I would never have been able to build complex optical imaging systems that are needed in my own experiment.

I would like to thank Dr. Cheng Shen, Dr. Jian Xu, Dr. Baptiste Blochet, and Federic de Goumoens for our wonderful collaborations and interesting brainstorming sessions. Thanks to Dr. Chi Shing (Antony) Chan, Dr. Joshua Brake, and Dr. Jaebum (Albert) Chung for our fruitful in-person discussions before the pandemic. Thanks to other members of the Biophotonics Lab, Mingshu Liang, Haowen Zhou, Max Huang, Zhenyu Dong, Siyuan Yin, Steven Lin, Shi Zhao, Dr. Oumeng Zhang and Dr. Simon Mahler, for the great lunchtime discussions. Thanks to Anne Sullivan, for her help in all kinds of lab and research related stuff.

Thanks to my friends, Xiaoqiao Chen, Yifan Chen, Yongzhao Guo, Yuchun Sun, Cheng Shen, Tian Xie, Zhiquan Yuan, Ding Zhong, Tianzhe Zheng, and Yang Zhang, for the great food we had, board game nights we spent together and many other fun activities we organized before, during, and after pandemic.

Thanks to my high school and lifelong friend, Jinbi Lu, for supporting me in my darkest days.

Thanks to my father, for his unconditional support in every single choice I made. Thanks to my mother, who is now in heaven, for her dearest love and support in the first two years of my Ph.D. study. Without them, I would not be who I am today. I am grateful for my parents who taught me to be patient if I found myself temporarily straying from the path towards my own goals when I was a child. Building new tools is always a long and challenging journey, and patience is essential in working out every detail to bring the new idea to life.

I went on to pursue a Ph.D. after I graduated from college with the eagerness to explore interesting problems and deepen my understanding in optics and science. I am happy that, despite all the hurdles I have experienced, I am still the same type of person who is hungry for new knowledge and enthusiastic about crazy yet intriguing ideas.

ABSTRACT

Optical imaging has gained popularity in life science, biomedical imaging, fundamental physics research, and various other fields due to its non-invasive nature. In a carefully designed optical instrument operating in an ideal environment, the resolution of the optical imaging system is defined by its numerical aperture. However, practical manufacturing issues and inaccurate lens models make it challenging to achieve high resolution across a large area. High magnification lenses introduce aberrations that degrade image quality, prompting the use of complex lens systems dedicated to mitigating such aberrations. Furthermore, when a scattering medium is introduced into the imaging system, image formation becomes infeasible as light follows a complicated trajectory. These challenges pose great obstacles to the use of optical imaging methods in various scenarios. This thesis primarily consists of two parts, one aims to deal with aberration and the other tries to solve scattering induced imaging problems.

In the first part of my thesis, I will discuss a technique called APIC (Angular Ptychographic Imaging with Closed-form method), which enables high-resolution imaging across a large field of view. To make APIC applicable in many non-ideal cases where aberrations (such as defocus) degrade image quality, we equip APIC with a closed-form aberration correction algorithm. We will demonstrate that APIC is unprecedentedly robust against aberrations and can retrieve high-resolution complex light fields using low magnification objectives.

In the second part, we move on to dealing with scattering induced imaging problems. To form images where a scattering medium is present, we first explore the application of ultrasound modulation in optical imaging. We show that, by using ultrasound, we can image a hidden object in a highly scattering medium with ultrasonic resolution. Although this technique helps obtain clear images in the presence of a scattering medium, its resolution is limited. We then demonstrate a method in addressing another scattering problem, namely the non-line-of-sight (NLOS) imaging problem. In a general NLOS problem, modulation mechanisms such as the aforementioned ultrasound modulation are infeasible. We demonstrate that light can be directly focused on the hidden target with an optical diffraction-limited resolution by exploring the properties of the hidden target itself. We will show that this active focusing method possess remarkably improved resolution compared to existing methods and is able to image objects with large reflectance differences.

PUBLISHED CONTENT AND CONTRIBUTIONS

1. Cao, R., Shen, C. & Yang, C. High-resolution, large field-of-view label-free imaging via aberration-corrected, closed-form complex field reconstruction. *arXiv preprint.* `http://arxiv.org/abs/2309.00755` (2023).
   R.C. conceived the idea, conducted the theoretical analysis, performed the simulations, participated in the experiments, wrote the reconstruction algorithm, and participated in the writing of the manuscript.

2. Cao, R., de Goumoens, F., Blochet, B., Xu, J. & Yang, C. High-resolution non-line-of-sight imaging employing active focusing. *Nature Photonics* **16,** 462–468. ISSN: 1749-4893. `https://www.nature.com/articles/s41566-022-01009-8` (2022).
   R.C. developed the idea, designed and developed the experimental protocol and setup, conducted the theoretical analysis, conducted the experiments, wrote the simulation and experiment codes, and participated in the writing of the manuscript.

3. Liu, Y., Cao, R., Xu, J., Ruan, H. & Yang, C. Imaging through highly scattering human skulls with ultrasound-modulated optical tomography. *Optics Letters* **45,** 2973–2976. ISSN: 1539-4794. `https://opg.optica.org/ol/abstract.cfm?uri=ol-45-11-2973` (2020).
   R.C. participated in the building of the imaging system and the writing of the controlling code, conducted the experiments, performed analysis on the raw data, and participated in the writing of the manuscript.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

NOMENCLATURE

Aberration. In imaging, aberration describes that some of the light spreads out in some regions other than the desired focused diffraction-limited spot. This leads to degradation in image quality such as a lower contrast and distortion. For coherent imaging system, it can be characterized by the phase of the coherent transfer function.

Coherent transfer function. A complex function $H(\mathbf{k}) \in \mathbb{C}$ lives in the spatial frequency domain, where $\mathbf{k} \in \mathbb{R}^2$ is the spatial frequency vector. It characterizes a coherent imaging system: its phase denotes the aberration and its amplitude determines the theoretical resolution of the imaging system.

Diffraction limit. The maximal theoretical resolution that can be achieved by an optical system. A diffraction-limited spot means the smallest focus spot that is achievable by an imaging system.

Field of view. The maximal area that can be imaged in an imaging system.

Objective. The lens on a microscope system that is placed near the sample.

Spatial frequency. When Fourier transform is applied to an image, we get a 2D transformed signal. Similar to Fourier transform applied to a time sequence, Fourier transform on a 2D image describes different frequency components of the image. As the transform acts in the spatial domain rather than the time domain, the frequency for a 2D image is called its spatial frequency.

Speckle. A type of grainy pattern with alternating bright and dark dots. It is typically observed when a laser is scattered, such as shining on a rough surface or passing through thick tissue.

Wavefront. A collection of all points sharing the same phase for a monochromatic light field.

Chapter 1

INTRODUCTION

Observations of new phenomena lead us to know more about our surrounding world and drive discovery of new physics. Measurement and sensing are fundamental tools for making these observations. Our eyes are also one type of such measurement tools that use light for the measurements. Due to its non-invasive and easy-to-use nature, light related measurements become a popular method in both our daily life as well as advance scientific researches.

Among different sorts of light related measuring apparatus, microscope is still one of the most important tools in biology, biomedical research, and beyond. Since microscope was invented, revolutionary developments in advanced modern microscopy shed light in imaging samples under complex scenarios with unprecedented resolution. These improvements are brought into reality with the help of development in various fields, including advances in optical design, sample preparation and labeling, light manipulation, post-processing, and various other areas. Such new development brings the possibility of imaging complex samples that we could never be able to by solely using conventional microscope objectives.

A core part in a typical microscopic imaging system is its objective lens, which ultimately determines the maximal resolution and field-of-view (FOV) that can be achieved in conventional microscopic imaging. However, in cases where a high-resolution, large FOV images is desired or clear images of sample in a scattering media is needed, conventional objectives fail to meet such demand.

In this chapter, I will briefly introduce some basic concepts with respect to light itself and an imaging system. With these established concepts, I will then introduce some existing problems in current imaging systems and methods I developed in my PhD in solving problems in optical imaging and improving current imaging techniques. The implementation of each methods will be discussed in detail in the following chapters.

1.1 Lens, spatial frequency, and resolution

Lens focuses a plane wave into a tight spot, which is called the focus and the plane contains for such focused spot is called the focal plane. Compared with the on-axis ray, a light ray that is parallel to the optical axis and is further away from the optical axis has a larger angle β after being focused by a lens, as shown in Fig. 1.1.

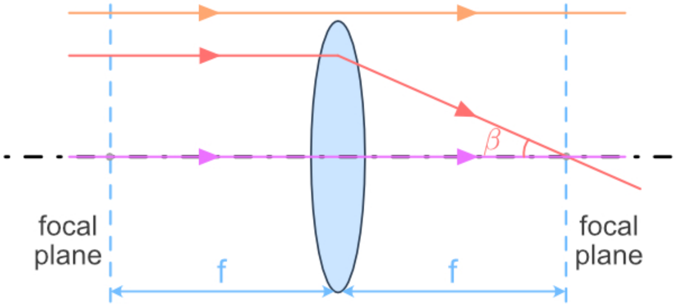

Figure 1.1: Focusing light using a lens. Here three different light rays are shown. The orange ray does not encounter the lens. The other two rays (red and purple) are focused by the lens.

Using Fourier optics, when a sample is placed at the focal plane of a thin lens, its E-field and the imaged E-field on the other focal plane can be reasonably treated as Fourier transform pairs [1]. Although such approximation is no longer a good one in some scenarios, one of such is lenses with an extreme curvature, we take this approximation to be a valid one in this thesis as most of the objectives used in reality can be treated as thin lenses. When a sample is placed at the focal plane of a lens, we call this focal plane the real space. Once this field is Fourier transformed by a lens, its Fourier transformed field sits on the other focal plane (which is called the Fourier plane). The space for the Fourier transformed signal is termed as the spatial frequency domain. To restore the light field, we can use another lens to perform the "inverse" Fourier transform (up to a constant and point reflection over the origin). We can see this point when we introduce Fourier transform in the following section. Performing Fourier transform and its inverse is the basic idea of a 4f system where two lenses are used for image capturing, as shown in Fig. 1.2.

We can see an ideal lens performs Fourier transform on a sample's field. However, due to limited size of a practical lens, not all light from the sample could pass through the lens: some of the light may not interact with the lens at all, such as the orange ray in Fig. 1.1. Thus, for a practical lens, the support (non-zero part) of its Fourier transformed field is confined: In the case of a point source on the sample plane, it will be converted to a "plane wave" when interacting with the lens and the width of the "plane wave" is limited.

For an ideal imaging system, we want an ideal point from the sample to be imaged

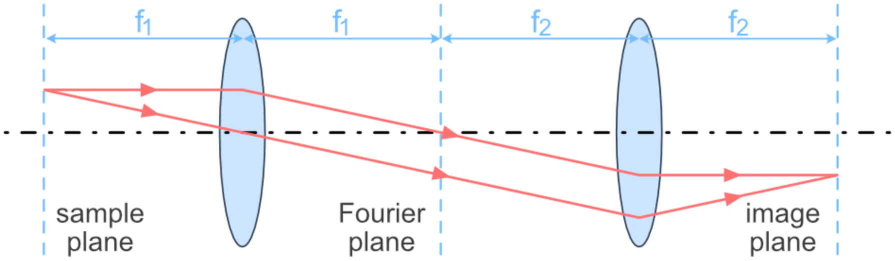

Figure 1.2: A typical 4f system. When the sample is placed at the front focal plane of the first lens and the distance between the two lenses equals to the sum of their focal lengths, the image of the sample is formed at the back focal plane of the second lens. The back focal plane of the first lens, which is exactly the front focal plane of the second lens, is called the Fourier plane as the Fourier transform of the light field lives in this plane.

onto an ideal point. This means that the sample's field is reproduced on the image plane without loss of the information. That is to say, we want the system to transmit the Fourier pairs without dropping information. Because a point (modeled by a Dirac delta function) has a Fourier transform which equals a non-zero constant in the entire domain, an ideal lens should have no constraint on the support of the Fourier transformed field. However, due to the practical size of a lens, some of the information was dropped in this process. In reality, thing is a bit more complicated than this simple example as the light is a form of wave and the maximal information can be carried by the light is limited as well. We will later introduce a wavelength dependent characterization for resolution of an imaging system.

Figure 1.3: Light diffraction when interacting with a grating.

To begin, we first give some intuitive examples on plane wave interacting with different gratings. If we have two types of grating, one with fine structure and the other with coarse structure, we shine a plane wave onto these two gratings. The diffracted light is shown in Fig. 1.3. Assume the grating has a period of d and the wavelength of the incident light is λ, the angle α between the first-order diffraction light and the incident light is given by

$$\sin(\alpha) = \lambda / d. \tag{1.1}$$

Thus, a finer grating results in a larger diffraction angle. This suggests that the high spatial frequency signal is associated with a larger angle. Based on this intuition, the metric we would use to quantify the resolution of an imaging system should characterize the angle.

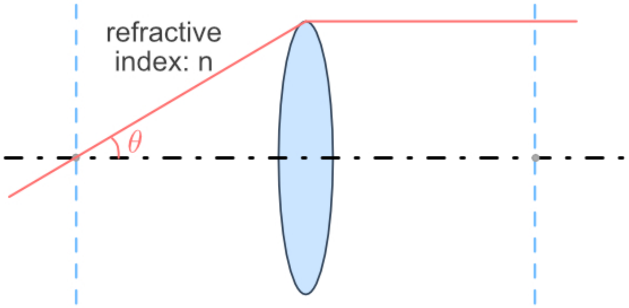

Figure 1.4: Definition of the numerical aperture of a lens.

As light with shorter wavelength can in theory carry more information than that of the longer wavelength, we should take the refractive index, which leads to changes in the wavelength, into consideration. We thus define the numerical aperture of a lens as

$$\mathrm{NA} := n \sin\theta, \tag{1.2}$$

where n is the refractive index of the surrounding medium, and θ is the maximal acceptance angle of a lens, as illustrated in Fig. 1.4. The resolution of a coherent imaging system, which can be calculated using diffraction theory, is given by

$$\text{Resolution} = \frac{\lambda}{n \sin\theta} = \frac{\lambda}{\mathrm{NA}}, \tag{1.3}$$

where λ is the wavelength of the incident light in vacuum. This formulae matches well with our intuition: a shorter wavelength in vacuum or a larger refractive index provides better resolution; and a larger acceptance angle lets more information to pass through the lens and preserves finer details of a sample.

1.2 Fourier transform and its properties

In this section, a brief introduction on Fourier transform is included. Some important properties of Fourier transform are presented without going into the derivation.

These properties are quite useful in understanding the effect of a lens in an imaging system. We will discuss the 2D Fourier transform as this thesis focuses on imaging one particular plane of a sample.

Definition 1.1 *Let the sample function be* $S(\mathbf{x})$ *and* $\mathcal{F}$ *be the Fourier transform operator. Let* $\hat{S}$ *be the Fourier transform of the original sample function. The Fourier transform of* $S(\mathbf{x})$ *is defined as*

$$\hat{S}(\mathbf{k}) = \left[\mathcal{F}(S)\right](\mathbf{k}) := \int S(\mathbf{x})e^{-i\mathbf{k}\cdot\mathbf{x}}d\mathbf{x}, \tag{1.4}$$

where $\mathbf{k} \in \mathbb{R}^2$ *denotes the spatial frequency,* $\mathbf{x} \in \mathbb{R}^2$ *denote the spatial location, and* i *is the unit imaginary number. The inverse Fourier transform* $\mathcal{F}^{-1}$ *is given by*

$$S(\mathbf{x}) = \left[\mathcal{F}^{-1}(\hat{S})\right](\mathbf{x}) := \frac{1}{2\pi}\int \hat{S}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}d\mathbf{k}. \tag{1.5}$$

From the definition of Fourier transform and the inverse Fourier transform, we can see these two are similar. Fourier transform has an extra minus sign on the exponent and does not carry a constant in front compared with the inverse Fourier transform. When we ignore the constant that is usually unimportant in imaging, we can also say a thin lens performs inverse Fourier transform by adapting the positive directions of the coordinate system. This explains why a 4f system can be applied for imaging purpose.

Fourier transform is extremely useful in optics as the light field from the sample is Fourier transformed when interacting with a thin lens of a microscope system. Note that the Fourier transform of a practical light field always exists because its energy is limited (its L2 norm is finite). Its inverse Fourier transform, as a consequence, always exists as well. Here, we will list some of the most important properties of Fourier transform.

Before we introduce its properties, we first define the convolution operator. Let $M(\mathbf{x})$ be another field, then the convolution of $S(\mathbf{x})$ and $M(\mathbf{x})$ is given by

$$[S * M](\mathbf{x}) := \int S(\mathbf{y}) * M(\mathbf{x}-\mathbf{y})d\mathbf{y}, \tag{1.6}$$

where $*$ stands for the convolution operator and $[S * M](\mathbf{x})$ means the convolution $S * M$ is evaluated at point $\mathbf{x}$.

Let $\hat{M}(\mathbf{k})$ be the Fourier transform of $M(\mathbf{x})$ and $a, b \in \mathbb{C}$ be some constants. Using the same notation used in the definition of Fourier transform, we have the following properties summarized in Table 1.1.

Function	Fourier transform	Remarks
$aS(\mathbf{x}) + bM(\mathbf{x})$	$a\hat{S}(\mathbf{k}) + b\hat{M}(\mathbf{k})$	linearity
$S(\mathbf{x} - \mathbf{x}_0)$	$\hat{S}(\mathbf{k})e^{-i\mathbf{k}\cdot\mathbf{x}_0}$	shift in real space
$S(\mathbf{x})e^{i\mathbf{k}_0\cdot\mathbf{x}}$	$\hat{S}(\mathbf{k} - \mathbf{k}_0)$	shift in spatial frequency space
$S(\mathbf{x}) * M(\mathbf{x})$	$\hat{S}(\mathbf{k})\hat{M}(\mathbf{k})$	convolution theorem
$S(\mathbf{x})M(\mathbf{x})$	$\frac{1}{2\pi}\hat{S}(\mathbf{k}) * \hat{M}(\mathbf{k})$	convolution theorem

Table 1.1: Properties of Fourier transform

One interesting example of Fourier transform is the Fourier transform of a Dirac delta function $\delta(\mathbf{x})$, which is defined as

$$\delta(\mathbf{x}) = 0, \forall \mathbf{x} \neq \mathbf{0} \text{ and } \int \delta(\mathbf{x}) d\mathbf{x} = 1. \tag{1.7}$$

We can see that this is basically the meaning of an ideal point in $\mathbb{R}^2$. We can further calculate its Fourier transform, which is

$$\hat{\delta}(\mathbf{k}) = \int \delta(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} d\mathbf{x} = 1. \tag{1.8}$$

For a shifted version of the delta function, namely $\delta(\mathbf{x} - \mathbf{x}_0)$, we can use Fourier transform's property to obtain the corresponding Fourier transform

$$\hat{\delta}_{-\mathbf{x}_0}(\mathbf{k}) := \mathcal{F}\big[\delta(\mathbf{x} - \mathbf{x}_0)\big](\mathbf{k}) = e^{-i\mathbf{k}\cdot\mathbf{x}_0}. \tag{1.9}$$

Let us see what happens to the resultant real space signal when we multiply the function $\hat{M}(\mathbf{k})$ with $\hat{\delta}_{-\mathbf{x}_0}(\mathbf{k})$ in the Fourier domain. This inverse Fourier transform of the multiplication gives

$$\Big(\mathcal{F}^{-1}\big[\hat{\delta}_{-\mathbf{x}_0}(\mathbf{k})\hat{M}(\mathbf{k})\big]\Big)(\mathbf{x}) = \Big(\mathcal{F}^{-1}\big[e^{-i\mathbf{k}\cdot\mathbf{x}_0}(\mathbf{k})\hat{M}(\mathbf{k})\big]\Big)(\mathbf{x}) = M(\mathbf{x} - \mathbf{x}_0). \tag{1.10}$$

Thus, the shifted delta function effectively translates $M(\mathbf{x})$ by the same amount when the two are multiplied in the spatial frequency domain. This leads to a powerful observation: it indicates that if an arbitrary point $\mathbf{x}$ on the sample, when interacting with a linear (imaging) system, is transformed into a particular pattern M that has a fixed offset with respect to the point $\mathbf{x}$, the system can be characterized by $\hat{M}$, the Fourier transform of this pattern.

1.3 Coherent transfer function and aberration

We know that a lens performs Fourier transform on a light field. Due to its limited size, the support of the Fourier transform is confined. This is one of the constraints of a practical lens in terms of the amplitude. So, how about the phase? Will a practical lens also introduce phase modulation to the Fourier transform of a light field? The answer is yes, we would also expect phase modulation due to the imperfections in the design and intrinsic material properties of a lens. In this section, we will introduce the coherent transfer function associated with an imaging system.

It is natural to assume that the amplitude and phase modulation induced by a practical lens is invariant with respect to the sample. A general modulation function can then be written as $H(\mathbf{k};\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^2$ denote the spatial location in real space (the image plane) and $\mathbf{k} \in \mathbb{R}^2$ denotes the spatial frequency in the Fourier domain (the Fourier plane). The phase of $H(\mathbf{k};\mathbf{x})$ denotes the associated phase modulation, and its amplitude is the associated amplitude modulation. Note that the argument $\mathbf{x}$ indicates that this function might vary with respect to the spatial location.

To do a further investigation on the properties on the phase and the amplitude of this modulation function, we first assume that the modulation is independent of the spatial coordinate. This is based on the idea that we expect an imaging system to map any given point on the sample plane to a fixed pattern, say another point, on the image plane (up to some translation depending on where the original point is). With this assumption, the general modulation function reduces to a complex function $H(\mathbf{k})$ independent of the spatial location $\mathbf{x}$. Based on our observation in section 1.2, the modulation function $H(\mathbf{k})$ should multiply with the sample's Fourier transform $\hat{S}(\mathbf{k})$. The resultant light field on the image plane $R(\mathbf{x})$ is then given by

$$R(\mathbf{x}) = \mathcal{F}^{-1}\left[H(\mathbf{k})\hat{S}(\mathbf{k})\right](\mathbf{x}). \tag{1.11}$$

From the above equation, we can see that an imaging system can be characterized by a simple function. We call this function the coherent transfer function (CTF). Although we can only use CTF to represent an imaging system given the modulation is invariant with respect to the spatial coordinate, it remains a good approximation if the modulation function varies slowly within a certain range, which is exactly the case for a typical imaging system.

For visible light regime, the frequency of light is much larger than the response of a normal camera. Thus, we cannot directly measure the light field. Instead, we measure the intensity I of the light field

$$I(\mathbf{x}) = |R(\mathbf{x})|^2 = R(\mathbf{x}) \cdot R^*(\mathbf{x}), \tag{1.12}$$

where $*$ on the superscript denotes the complex conjugate of a complex number. Thus, in our real measurement, the phase of the light field is lost and its intensity (amplitude squared) is recorded by the camera.

We discussed that high spatial frequency is associated with large diffraction angle (Fig. 1.3), thus, the high spatial frequency part may pass through the imaging system or being blocked due to the finite lens aperture. With respect to the amplitude of the CTF, this suggests the amplitude modulation is unity if we assume there is no attenuation in the light transmitting through the system, and zero when light is blocked. This turns out to be correct for most microscope system.

For the phase modulation, things are a bit more involving. Different from the attenuation that is much easier to minimize in reality, phase is way more sensitive as it depends on the wavelength of the light, which is only several hundred nanometers for visible light. For an ideal thin lens, the phase modulation should be zero as ideal lens performs Fourier transforms without altering the field. However, the phase modulation, termed aberration, always exists for a practical imaging system. This can be caused by not only the imperfections of manufacturing these lenses but also the inaccurate lens model (the paraxial approximation for example). To characterize this imperfection, we can use some basis to represent the phase modulation function such as the sinusoidal functions used in Fourier transform. However, these basis might not be efficient in a sense that we might need a great number of them for a reasonable representation.

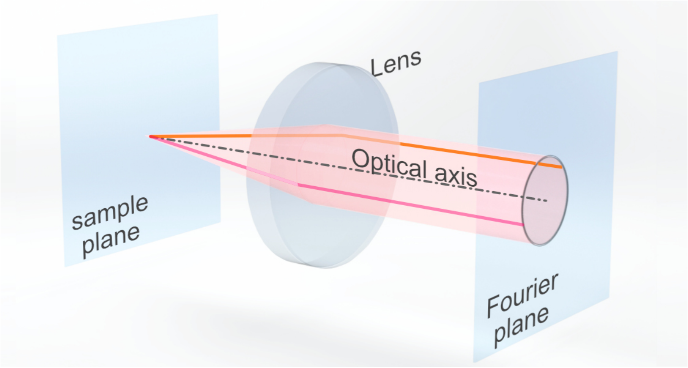

Figure 1.5: Two light beams that are symmetric with respect to the optical axis. Due to the symmetry of our real space, the orange and the magenta beams should have the same phase change when the sample plane moves in the axial direction.

Let us first focus on one common aberration, namely defocusing, where the structures in the image plane changes with respect to the defocus distance. In this case, defocus distance is the only varying parameter and we want to characterize this change using one variable instead of changing a large number of variables. For defocus, we may assume the phase modulation should be rotational invariant when using a thin lens: If we have a ring that is perpendicular to the optical axis and the center of this ring coincides with the optical axis, the phase modulation induced by defocus is the same along this ring (Fig. 1.5). Exploring such properties, these might exist a basis can be more efficient in representing the aberration.

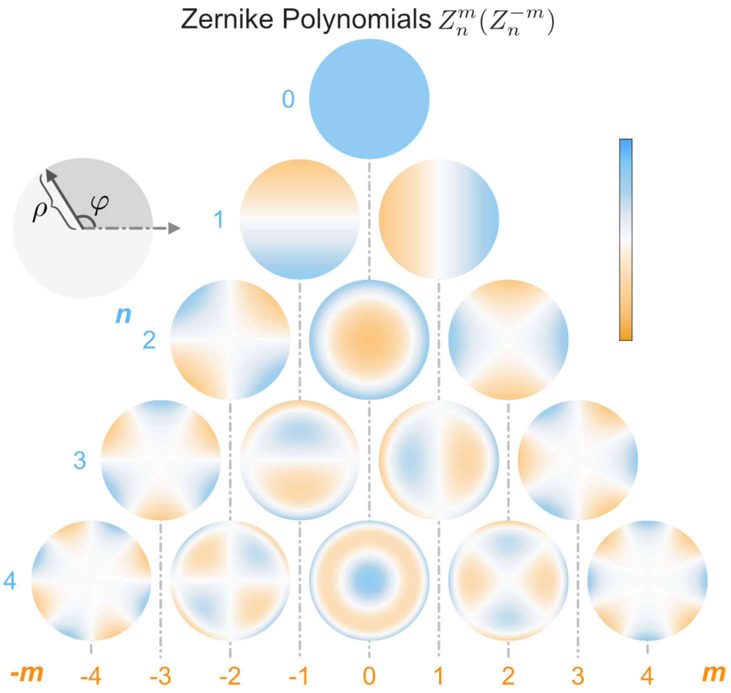

Figure 1.6: Visualization of the first 15 Zernike polynomials. The rows indicate different n and the columns indicate different m $(-m)$.

An efficient basis for characterizing the aberration does exist, which is called the Zernike polynomials. The Zernike polynomials consist an orthogonal basis on the unit disk, and the first tens of polynomials can already well represent the aberration of an typical imaging system in practice. Zernike polynomials can be factorized into the radial part and the azimuthal angle part: let m and n be non-negative integers subject to $n \geq m \geq 0$, we define the radial function $\mathcal{R}_n^m$ as

$$\mathcal{R}_n^m(\rho) = \begin{cases} \sum_{j=0}^{\frac{n-m}{2}} \frac{(-1)^j\,(n-j)!}{j!\left(\frac{n+m}{2}-j\right)!\left(\frac{n-m}{2}-j\right)!}\,\rho^{n-2j}, & n-m \text{ is even,} \\ 0, & n-m \text{ is odd.} \end{cases} \tag{1.13}$$

The Zernike polynomials are given by

$$Z_n^m(\rho,\varphi) = \mathcal{R}_n^m(\rho)\cos(m\varphi) \quad \text{and}$$
$$Z_n^{-m}(\rho,\varphi) = \mathcal{R}_n^m(\rho)\sin(m\varphi), \tag{1.14}$$

where Z_n^m denotes the even polynomials and Z_n^{-m} denotes the odd polynomials. The visualization of the first few Zernike polynomials can be found in Fig. 1.6.

The first three Zernike terms are typically not of interest as they correspond to a global phase offset or a shift in the absolute focus position. These do not change the structural properties of the imaged field. The fifth term, Z_2^0, represent the aberration associated with the defocus of a sample. It turns to be rotational invariant which matches well with our previous argument.

We note that due to the refractive index's wavelength dependency, light of a different frequency might have a different aberration. Thus, to compensate the aberration in multi-spectral imaging where lights of multiple wavelengths are used, we need to consider the aberration for each of the wavelength used in experiments.

1.4 Wavefront, phasor, and scattering

We learnt from optics course that the wavefront of a monochromatic light field describes a collection of all points which share the same phase. Important examples of wavefront include plane wave and spherical wave, as shown in Fig 1.7. The solid red lines denote the points having the same phase and visualize the wavefront of light.

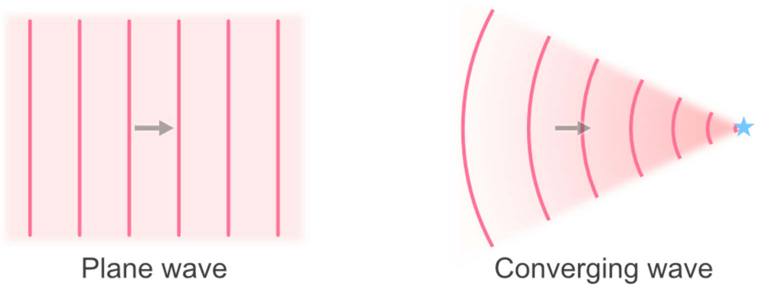

Figure 1.7: Wavefront of a plane wave and a converging spherical wave.

A typical coherent imaging system replicates the sample's wavefront on the detector. However, when the sample is placed inside a scattering medium or sample is thick and we want to image one plane in the middle of the thick sample, the sample associated wavefront might be distorted due to scattering. Besides scattering, other types of light-matter interactions such as absorption also occur simultaneously. In most biological samples, however, the effect of absorption is small compared scattering. We will thus focus on the scattering induced problem in imaging and ignore many other mild phenomena.

With respect to the wavefront of light, scattering adds random phase modulation to the forward scattered field of the incident wavefront. Due to the randomness and stochastic nature of scattering, the resultant phase modulation does not exhibit symmetry. So, it is no longer efficient to characterize the random phase modulation using Zernike polynomial. Moreover, we cannot assume the phase variation is invariant with respect to different points on the sample as light from different points interacts with different part of the scattering medium. Based on these reasons, we will no longer use CTF in the presence of highly scattering sample and will instead focus on its wavefront.

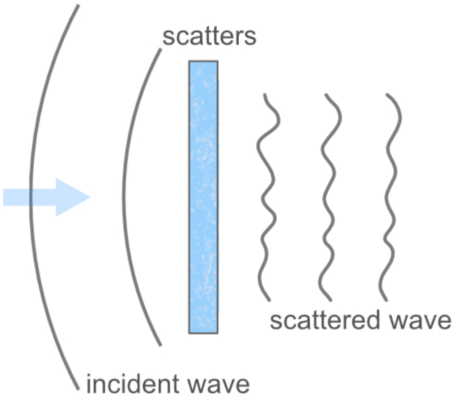

Figure 1.8: Spherical wave passing through a scattering medium.

For the convenience of analysis, we first introduce the phasor of a complex number. Although light has a fast-changing electric field in time, two points in a monochromatic light field nonetheless have a time-invariant phase difference in equilibrium. We thus define the phasor of light with respect to a global reference. With this definition, the phasor of the light becomes time-invariant for fixed samples.

Assume the complex field at a point x has amplitude A_x and relative phase ϕ_x with respect to the reference, we write the field E at point x as

$$E(x) = A_x e^{i\phi_x}. \tag{1.15}$$

In the complex plane, we can use an arrow to represent this complex number. The length of the arrow denotes the amplitude and the angle in between the arrow and the positive real axis denote the phase of the complex number. We call this arrow as the phasor, which is shown in Fig. 1.9.

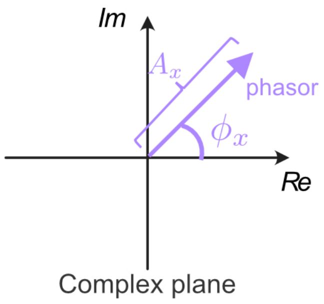

Figure 1.9: Phasor of a complex number $A_x e^{i\phi_x}$. “Im” denotes the imaginary part of a complex number and “Re” describes its real part. The length of the arrow denotes the amplitude of a complex number and the angle of the arrow denotes the phase of this complex number.

Before we talk about scattering, we will talk about the interference of light because it is essential to understand the scattered light field. When a complex light field interact with another light field, its resultant amplitude will be the superposition of the two fields. Let us assume the two fields are given by $E_1(x) = A_{1,x}e^{i\phi_{1,x}}$ and $E_2(x) = A_{2,x}e^{i\phi_{2,x}}$, the superposition of the two is

$$E_{\text{superposition}} = E_1(x) + E_2(x) = A_{1,x}e^{i\phi_{1,x}} + A_{2,x}e^{i\phi_{2,x}}. \tag{1.16}$$

A detector measures the intensity of this superpositioned field, which yields

$$\begin{aligned} I_{\text{superposition}} &= |E_{\text{superposition}}|^2 = \left|A_{1,x}e^{i\phi_{1,x}} + A_{2,x}e^{i\phi_{2,x}}\right|^2 \\ &= A_{1,x}^2 + A_{2,x}^2 + 2A_{1,x}A_{2,x}\cos(\phi_{1,x} - \phi_{2,x}). \end{aligned} \tag{1.17}$$

In terms of the phasor, the above two equations can be visualized by Fig 1.10.

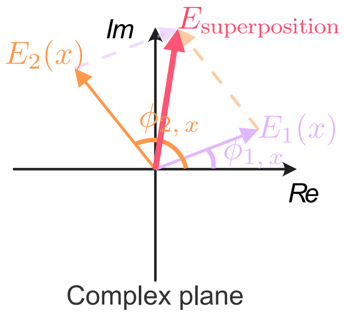

Figure 1.10: Superposition of two fields.

Based on Huygens principle, we can treat each point of the new wavefront as a point source. Let us consider a toy case where we put a thin scattering sample and see its resultant field. In this case, the wavefront of the field exiting the scattering medium is reshaped. Points on this wavefront in Fig. 1.11 are then treated as multiple point sources with different amplitudes. The field at a further distance is the superposition of these spherical waves originated from the corresponding point sources. Due to the random wavefront where the Huygens principle is applied to, the propagated phasors should also have random phases. Thus, the resultant field is a superposition of multiple random phasors in the presence of scattering.

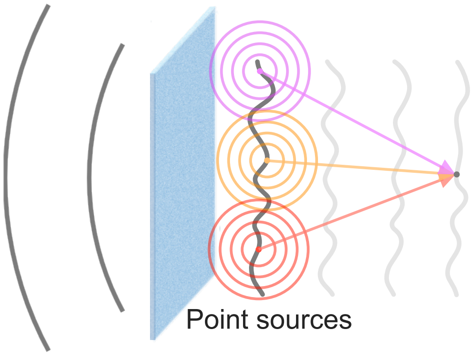

Figure 1.11: The Huygens principle.

The speckle, a type of granular pattern associated with scattering, is essentially a resultant of a random phasor sum. Owing to the randomness of the orientation of each phasor, their resultant phasor varies in its length. Thus, the captured intensity pattern exhibits alternating bright and dark grains. Three examples of the random phasor sums are illustrated in Fig. 1.12.

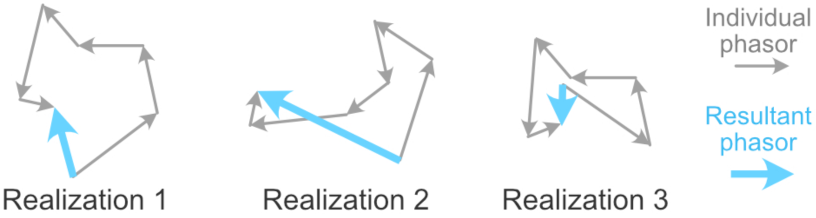

Figure 1.12: Examples of random phasor sums. Here, the lengths of individual phasors are fixed and only their angles are randomized in each realization.

As we can see from Fig. 1.12, the length of the resultant phasor varies considerably. For coherent light, the randomness in phase results in spatially random reconstructive and destructive interference of light, which ultimately contributes to the random spatially varying intensities. Figure 1.13 shows a representative image of a speckle pattern. Because of the random phase modulation, one point on the image side will not be imaged onto another point when a conventional imaging system is used. Thus, the presence of scattering medium in an imaging system greatly obstructs the restoration of the useful signal.

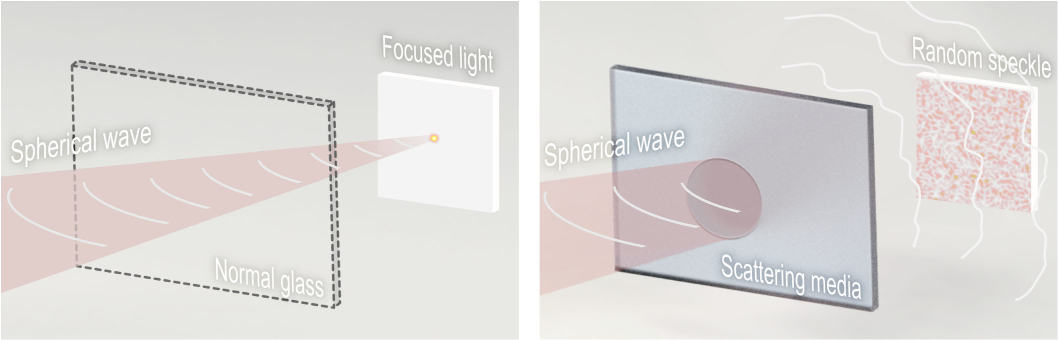

Figure 1.13: Images of a speckle pattern and a normal focus. When a scattering medium is inserted, the original wavefront is randomized. A speckle pattern thus shows up on the screen.

1.5 Thesis outline

This thesis focuses on two main aspects in imaging: dealing with scattering associated imaging problem and imperfect imaging system (aberration and lost of phase information) related problem.

In the second chapter, I will discuss a method we developed to simultaneously extract the lost phase information and correct for aberration using intensity measurements. By using tilted illuminations, the resolution of an imaging system can be pushed above the conventional NA limit of an objective. We will establish a closed-form phase retrieval framework to reconstruct the sample's complex light field using intensity measurements from a typical camera, without the use of interferometry.

In the third chapter, method based on light-ultrasound interaction was employed to circumvent the scattering induced imaging problem. We show that by using ultrasound as a gating mechanism, we can improve the imaging quality to that defined by the ultrasound, while maintaining the ability of detecting light absorption properties.

In the fourth chapter, we focus on a special and important scattering related scenario, the non-line-of-sight imaging. For the non-line-of-sight imaging, the direct line of sight is blocked. However, light interacting with the sample of interest can be scattered, say by a wall, and then enter our imaging system. Here, the sample's information is still in presence in the captured light. However, due to scattering, it is extremely hard to resort the sample's information. We showed that, with a proper design, we can replace the point source with an arbitrary sample of interest to do wavefront shaping, which compensates for the scattering induced random phase modulation.

We note that the symbols might differ from chapter to chapter. For the variables appeared in each chapter, please find their definitions in the corresponding chapter.

Chapter 2

A CLOSED-FORM, HIGH-RESOLUTION, LABEL-FREE IMAGING METHOD

In this chapter, we present an analytical method that can simultaneously correct aberration and reconstruct sample's complex light field using non-interferometric intensity measurements. We will first discuss some previously developed imaging methods that can be used to reconstruct the complex light field from an intensity measurement. Existing issues with those methods are discussed. To solve such issues, we are motivated to build a method that is more robust against aberration and more reliable in terms of the complex field reconstruction.

2.1 Background on the label-free, complex field reconstruction methods

To acquire the complex light field, one straightforward method is to use a reference beam to obtain an interferogram. In one of such techniques, called off-axis holography, an interferometry is built and a tilted reference beam is used to interfere with the sample beam. With the use of interferometry, the complex field can be easily reconstructed from pure intensity measurement with a guaranteed fidelity in theory. However, the use of an additional reference arm requires using lasers with good temporal coherence. In addition, mechanical shifts and air turbulence can be a problem as they disturb the phase of the light. Those can lead to inaccurate complex field reconstructions in practice.

There is also non-interferometric way of reconstructing the complex field from the intensity measurements. Phase retrieval, a type of computational method for finding the complex field use the intensity image, is the most promising way. In those computational methods, phase retrieval algorithms such as Gerchberg–Saxton (GS) algorithm [2] is used to reconstruct the missing phase information. One typical paradigm of the phase retrieval algorithm is shown in Fig. 2.1. This algorithm conduct alternating Fourier transform and inverse Fourier transform and poses different constraints in the real space and the spatial frequency space until we find the constraints are satisfied or the error is acceptable. Assume the aberration of the system is minimal, the NA-limited sample's field can then be retrieved using the GS algorithm.

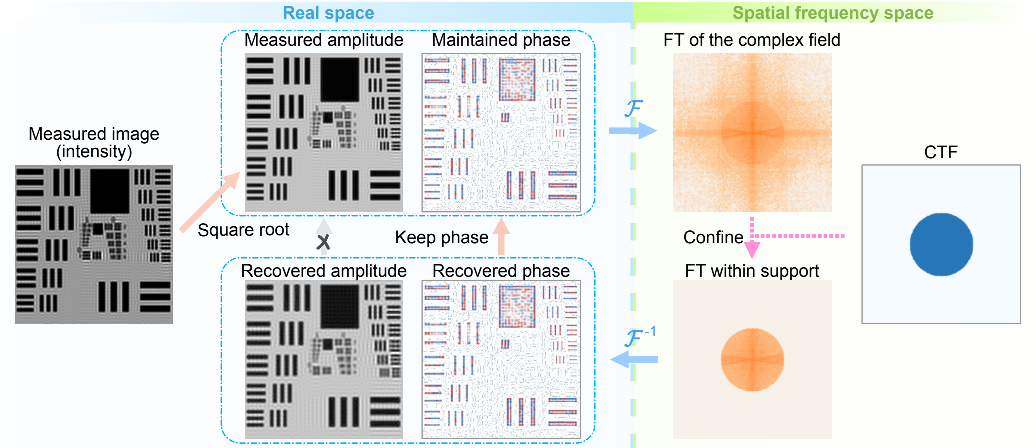

Figure 2.1: Gerchberg–Saxton algorithm. We first take the square root of the intensity measurement and treated its outcome as the correct amplitude of the complex field. The Gerchberg–Saxton algorithm conduct Fourier transform and inverse Fourier transform in an iterative way. In real space, the amplitude of the reconstructed complex field is replaced by the real measurement, and the phase is maintained. In the spatial frequency domain, we confine the Fourier transform of the reconstructed complex field as we know that only places covered by the CTF can have non-zero signal.

Although the GS algorithm enables recovering a complex field from the intensity measurement, it nonetheless needs to minimize the imaging system's aberration. Moreover, the resolution of the system is limited by the NA of the imaging system. In order to correct the aberration in an imaging system and obtain images with high resolution and large FOV, some other methods are desired. There is one special complex field reconstruction technique, called Fourier ptychographic microscopy (FPM), that leverages the power of computation to provide high-resolution and aberration correction abilities to low NA objectives [3, 4]. FPM operates by collecting a series of low-resolution images under tilted illumination and applies phase retrieval algorithm to reconstructs sample's high spatial-frequency features and optical aberration, resulting in high-resolution aberration-free imaging that preserves the inherently large FOV associated with the low numerical aperture objectives. In Section 1.1, we briefly discussed how high spatial frequency component is connected with the diffraction angle (Fig. 1.3). Here, we will discuss about the tilted illumination.

Consider a case where we shine a plane wave on a grating with a normal incidence angle, we expect that the diffracted light to be symmetric with respect to the grating, as shown on the left side of Fig 2.2. When the incident light is tilted, the diffracted light should also tilt towards the same direction. This means the light with large diffraction angle that cannot be collected by the objective might then enter the imaging system. This means that light that carries the high spatial frequency information can now be detected by the imaging system, while it is typically lost in the conventional microscopy.

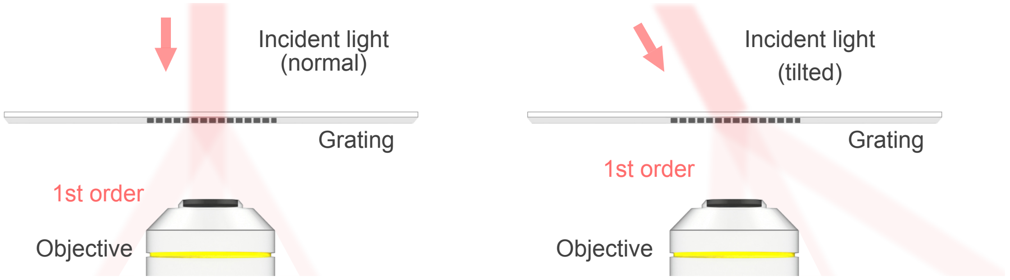

Figure 2.2: Diffracted light from a grating with different incidence angles. For normal incidence, we expect the diffracted light to be symmetric. For tilted light, the light should tilt along with the incident light. Under normal incidence, the diffracted light cannot be collected by the objective. When the incident light tilts, we see that the diffracted light enters the imaging system.

From the point of view of the spectrum, the tilted illumination brings the high spatial frequency spectrum into the detected range. Thus, using tilted illumination, we can in theory achieve a resolution that is beyond the NA-dependent limit. This principle is illustrated in Fig. 2.3, which is exactly the mechanism that enables FPM in achieving high-resolution images using low NA objectives.

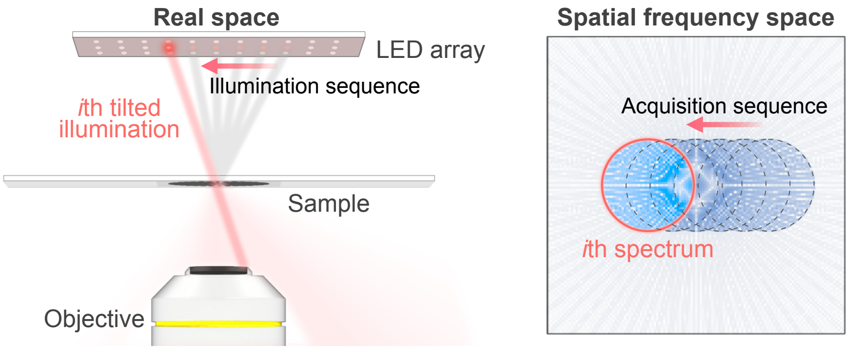

Figure 2.3: Principle of Fourier ptychographic microscopy (FPM). When using a series of tilted illuminations, we effectively translate the CTF to different place in the spatial frequency domain. Thus, different regions of the spectrum are sampled using different illumination angles.

To obtain the spectrum in each measurement, the complex field must be reconstructed. In order to achieve this, FPM utilizes phase retrieval algorithm for the complex field reconstruction. The reconstructed spectrums are then stitched together, which in turn gives us an improved resolution. To extract the aberration of the imaging system, FPM treats the aberration as another "complex field" and reconstructs it along with the complex sample field [4, 5]. This is feasible as the aberration and the sample's spectrum behave differently in the imaging process, as depicted in Eq. 2.6. By taking the difference into consideration, the aberration and the sample's spectrum can be isolated from each other. Detailed derivation can be found in previous works of FPM [5] and we will not go into detail in this thesis. Due to these attractive traits, FPM has found diverse applications in quantitative phase imaging, aberration metrology, digital pathology, and other fields [4, 6].

2.2 Principle of APIC, a closed-form field reconstruction method

Although FPM is an important advancement in label-free microscopy, its essential iterative reconstruction algorithm poses several challenges. First and foremost, the iterative reconstruction of FPM is a non-convex optimization process, which means that it is not guaranteed to converge onto the actual solution [2–5, 7–9]. In practice, the algorithm executes alternating projections between real space and spatial frequency space until certain conditions are met, such as its loss function decreasing rate reaches a lower bound, or the execution reaches the allowed maximum iteration number, or the algorithm satisfies other pre-defined metric thresholds [2–4, 7, 8, 10, 11]. As a result, FPM does not guarantee that the global optimal solution is ever reached. This is problematic for exacting applications, such as digital pathology, where even small errors in the image are not tolerable. Furthermore, the joint optimization of aberration and sample spectrum can fail when the system's aberrations are sufficiently severe — leading to poor reconstructions [12]. The iterative nature of FPM reconstruction algorithm has prompted researchers to adapt machine learning concepts to its implementation, in pursuit of computational load reduction, artifact abatement and aberration correction [13–16]. These in turn lead to other problems, such as contextual sensitivity and potentially greater drift away from the global optimal solution. It is worth considering at this juncture whether it is possible to develop a closed-form solution to this class of computational imaging problems, so that all these challenges can be more effectively addressed.

Recent studies have shown that the complex field can be non-iteratively reconstructed in one specific varied illumination microscopy scenario by matching the illumination angle to the objective's maximal acceptance angle (the NA-matching angle) and exploiting the signal analyticity, for example through spatial-domain

Kramers-Kronig imaging [17–19]. These findings are important and impactful as they eliminate the need for an iterative reconstruction framework and do not require a human-engineered convergence criterion. However, it is worth noting that this approach does not possess the capability to correct hybrid aberrations nor provide great resolution enhancement beyond the diffraction limit of the objective NA. As such, FPM remains a more appealing choice in various scenarios.

In this chapter, we present a novel analytical method, termed Angular Ptychographic Imaging with Closed-form method (APIC), that weds the strengths of both methods. APIC builds on complex field reconstruction using Kramers-Kronig relations and employs analytical techniques to retrieve aberration and reconstruct the darkfield associated high spatial frequency spectrum. By using NA-matching and darkfield measurements, APIC is capable of retrieving high-resolution, aberration-free complex fields when a low magnification, large FOV objective is used for data acquisition. From both simulations and experiments, APIC demonstrates unprecedented robustness against aberrations while FPM drastically fails. Due to its analytical nature, APIC is inherently insensitive to optimization parameters and offers a guaranteed analytical complex field solution. We additionally show that APIC can perform better than FPM when subjected to the same constraint on input data size, as it does not require an overly large data redundancy needed by FPM for a good convergence. By incorporating darkfield measurements, APIC effectively achieves the same theoretical resolution enhancement as FPM. The source code of APIC is available online (github.com/rzcao/APIC-analytical-complex-field-reconstruction). We believe APIC represents an impactful step forward in the field of computational imaging.

APIC collects both NA-matching and darkfield intensity measurements for high-resolution reconstruction. Its reconstruction process begins by analytically solving for the sample's spatial frequency spectrum and aberration with the NA-matching measurements. Then, the darkfield measurements are used to extend the sample's spatial frequency spectrum to greatly enhance the resolution of a NA-limited imaging system. The system setup, data acquisition process and its reconstruction flowchart are illustrated in Fig. 2.4. In APIC's data acquisition step, the LEDs whose illumination angles match up with the maximal acceptance angle of the imaging system are sequentially lit. The measurements under these NA-matching angle illuminations constitute the NA-matching measurements of APIC. LEDs whose illumination angles are greater than the acceptance angle are then successively lit for acquiring darkfield measurements. In the following sections, we use the word “spectrum” as shorthand for spatial frequency spectrum (the Fourier transform of the sample’s complex field). We note that the spectrum is different from the Fourier transform of an acquired image, which is the Fourier transform of a purely intensity measurement.

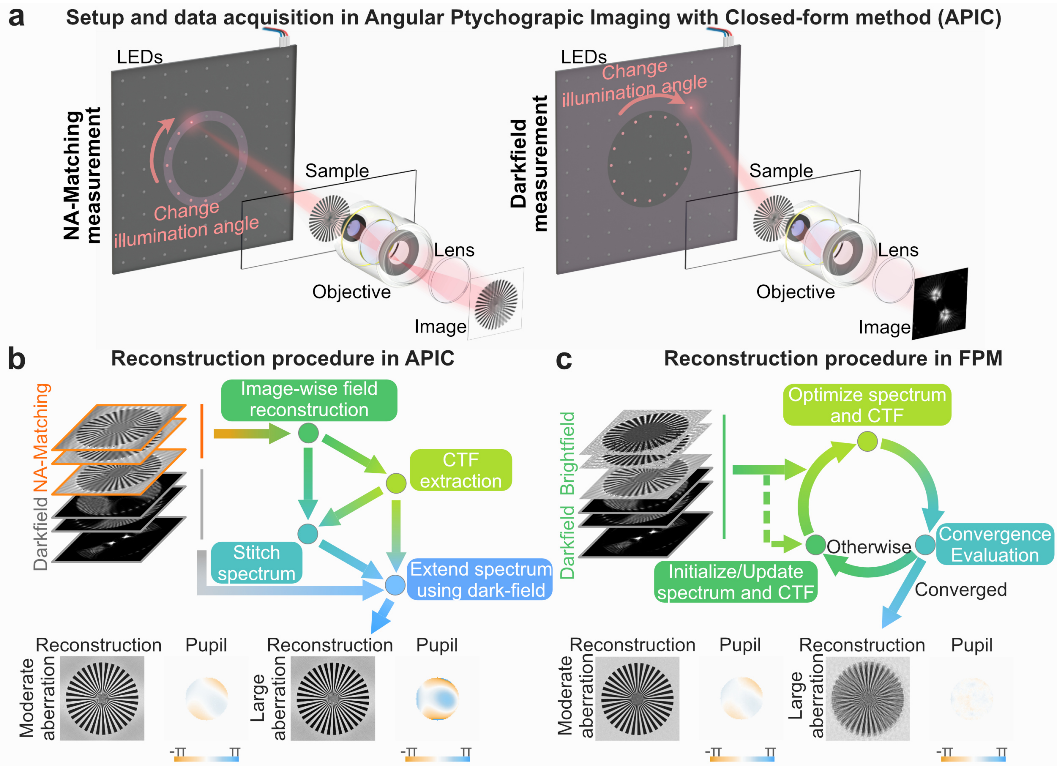

Figure 2.4: Concept of APIC and its difference compared with FPM. **a**, Setup of APIC. The LEDs whose illumination angle matches up with the numerical aperture (NA) of the objective are lit sequentially to obtain the NA-matching measurements. Then, the LEDs whose illumination angle is larger than the objective’s receiving angle are successively lit for the darkfield measurements. **b**, Reconstruction process of APIC. Once the aberration is extracted, it is used to correct aberration in the image-wise field reconstruction. The aberration-corrected spectrums are then stitched together and serves as a prior knowledge in the spectrum extension. Using darkfield measurements, the spectrum is furtherly extended to obtain a high-resolution, aberration-free reconstruction. **c**, Reconstruction process of Fourier ptychographic microscopy (FPM). FPM iteratively updates the spectrum and the aberration to minimize the differences in the measurement and reconstruction output. This iterative process is terminated upon convergence to obtain the spectrum and coherent transfer function (CTF) estimate. Pupil in the figure denotes the reconstructed aberrations.

APIC operates by first reconstructing the complex field corresponding to NA-matching measurements using Kramers-Kronig relations. These measurements are taken with LED illumination angles that match with the objective's maximal receiving angle. For a realistic imaging system, aberrations are inevitably superimposed on the spectrums' phase. To extract the objective's aberrations, we focus on the overlapping region in their spectrums (the overlap of two translated CTFs, as shown on the left side of Fig. 2.5). As the sample dependent phases are identical in the overlapped region of the two spectrums (note we assume a thin sample is used), subtracting their phases cancels out the sample dependent phase term, leaving only the phase differences between different parts of the pupil function (see Fig. 2.9 for more information). Consequently, the overlapping regions give us a linear equation with respect to the aberration term. By solving this linear equation, the aberration of the imaging system can be extracted, which can then in turn be used to correct the original reconstructed spectrums. The corrected spectrums are then stitched together to obtain an aberration-free, two-fold resolution-enhanced sample image.

We can then extend the spectrum by using the darkfield measurements. In this step, the reconstruction spectrum and the aberration obtained in the first step serve as the a priori knowledge. The step-by-step reconstruction operates in the following way. We choose a measurement whose spectrum is closest to the known spectrum (say, the ith measurement) and crop out the known spectrum based on what is sampled in this measurement, as shown in Fig. 2.5. This cropped spectrum, however, only contains part of the information of the ith measurement. Our goal is to recover the unknown spectrum so that it can be filled in for spectrum spanning.

We can see that the Fourier transform of our ith intensity measurement consists of cross-correlation of the known and unknown spectrum and their autocorrelations. In the following, we show that by using the known spectrum, we can construct a linear equation with respect to the unknown spectrum, which can be analytically solved.

First, the autocorrelation of the known part is calculated and subtracted from the measurement. After subtraction, the autocorrelation of the unknown part and the cross-correlations are left. One important observation is that these parts are not fully coincided in the spatial frequency domain (Fig. 2.5). As such, we can focus on the non-overlap region where the cross-correlation solely contributes to the signal.

We can then construct a linear equation with respect to the unknown spectrum. When calculating the cross-correlation, one of the signals is shifted and gets multiplied with another signal. The correlation coefficient is the summation of this product. Assuming one of the two signal is known, we can essentially use the known signal

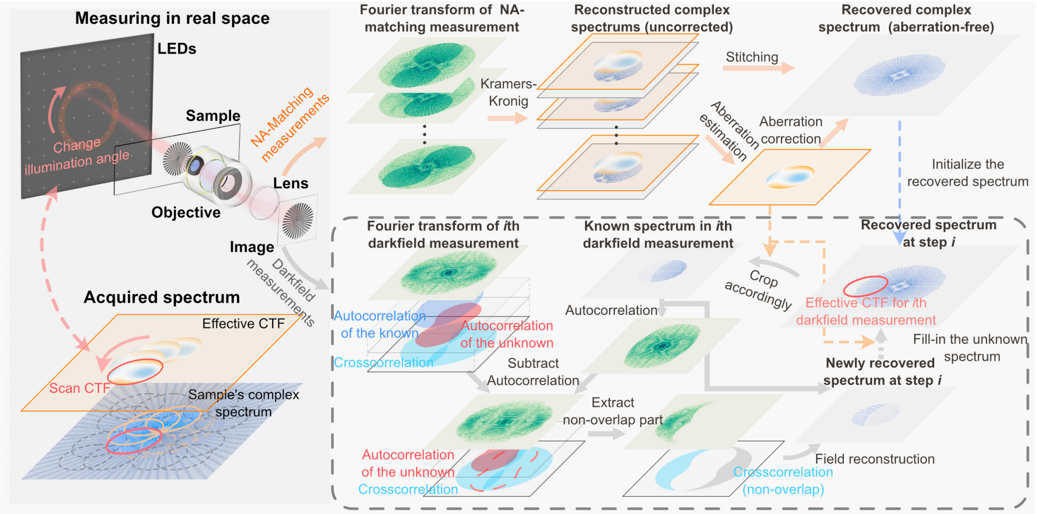

Figure 2.5: Reconstruction pipeline for APIC. By changing the illuminating angle, we effectively shift the CTF to different positions in the spatial frequency domain, and samples different regions of sample's spectrum. For measurements under NA-matching angle illumination, we first use Kramers-Kronig relation to recover the corresponding spectrums. The phase differences of two spectrums with overlaps in their sampled spectrum are used to extract the imaging system's aberration. Then, the image-wise reconstructed spectrums are corrected for aberration and get stitched, which forms our prior knowledge in reconstruction process involving darkfield measurements. To extend the spectrum using darkfield measurement, the known spectrum in the ith measurement is used to isolate cross-correlation from other autocorrelation terms. By solving a linear equation involving the isolated cross-correlation, the unknown spectrum can be analytically obtained. We then use the newly reconstructed spectrum to extend the original reconstructed spectrum.

as the weights and calculate the summation of a weighed version of the other signal. This is a linear operation. Thus, applying the known spectrum, we can construct a linear operator that takes the unknown spectrum and produces this cross-correlation. By extracting the non-overlapping part of the cross-correlation term, we can form and analytically solve a linear equation with respect to the unknown spectrum. That is, we obtain the closed-form solution of the unknown spectrum by solving this equation.

For a practical imaging system, we need to consider its aberration in the imaging process. To match up with the measurement, the recovered aberration is initially introduced to the cropped out known spectrum. After recovering the unknown spectrum, the aberration gets corrected, and this corrected spectrum is filled back into the reconstructed spectrum. This process stops when all darkfield measurement has been reconstructed. The detailed derivation of the aforementioned analytical complex field reconstruction and aberration extraction methods can be found in Section 2.6.

Once the above steps are completed, we obtain a high-resolution and aberration-free sample image. The theoretical optical resolution of APIC is determined by the sum of the illumination NA and the objective NA, which is identical to FPM's NA-resolution formulae [3]. We note that FPM requires an iterative process to recover the spectrum and is sensitive to the choice of the optimization parameters. On the other hand, APIC analytically recovers the actual spectrum. This direct and efficient approach sets APIC apart from FPM, offering a more straightforward and robust spectrum recovery process. In the following section, we will report on our experimental demonstration that APIC is computationally efficient and achieves consistent high-quality complex field reconstructions even under large aberrations, whereas FPM struggles due to the increased complexity in its optimization problem.

2.3 Experiment results of APIC

We used a low magnification objective (10x magnification, NA 0.25, Olympus) for data acquisition. A LED ring (Neopixel ring 16, Adafruit) glued onto a LED array (32x32 RGB LED Matrix, Adafruit) served as the illumination unit. The two LED clusters were mounted on a motorized stage for position and height adjustment, and they were individually controlled by two Arduino boards (Arduino Uno, Arduino). In the acquisition process, we lit up one LED at a time, and simultaneously triggered the camera (Prosilica GT6400, Allied Vision) to capture an image when the LED was on. This process continued until all usable LEDs was lit once. We then performed reconstruction using both APIC and FPM.

In our first experiment, we imaged a Siemens star target and chose to acquire a small dataset to perform reconstruction using APIC and FPM. The dataset acquired in this experiment consisted of 9 brightfield measurements, 8 NA-matching measurements and 28 darkfield measurements. We note that there are works that apply multiplexed illumination scheme to reduce the number of the measurements in FPM [7, 8], these methods are not as reliable as the conventional FPM data acquisition scheme. Thus, we only focus on the more reliable acquisition scheme in this study. The nominal scanning pupil overlap rate was approximately 65%. In our experiments, the second order Gauss-Newton FPM reconstruction algorithm was applied for reconstruction as it was found to be the most robust FPM reconstruction algorithm [9]. We also note that we used 6 sets of parameters in the reconstruction of FPM and chose the best result, as the reconstruction quality of FPM heavily depends on its parameters. Some representative FPM results are also shown in Fig. 2.6, which confirms such parameter dependency. On contrary, the faithfulness and correctness are guaranteed in APIC, benefiting from its new analytical phase retrieval framework. We found that APIC was able to render the correct complex field while FPM failed, as shown in Fig. 2.6.

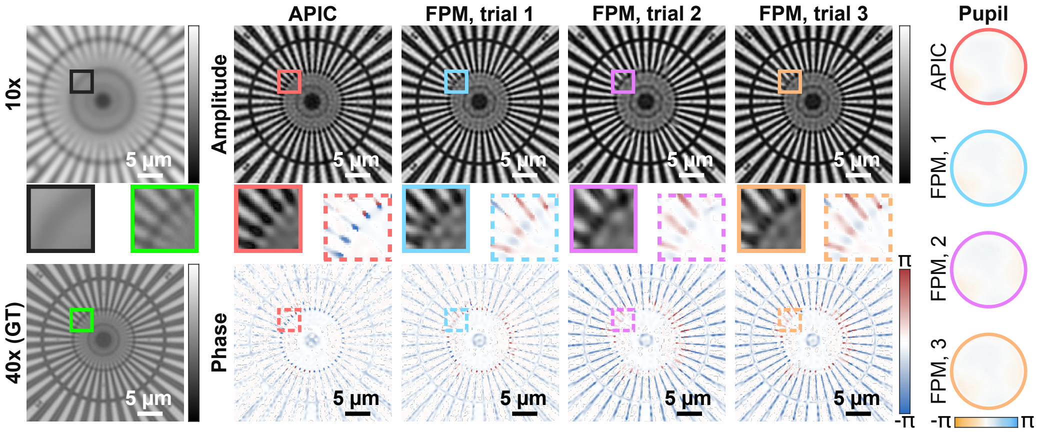

Figure 2.6: APIC and FPM results using a small number of measurements. For comparison, we also acquired a ground truth (GT) image, which was imaged under a high-NA, 40x objective. The Kohler illuminated image under the same 10x objective used in APIC's data acquisition is shown on the upper left. The reconstructed pupils are shown on the right side of the figure. For FPM reconstructions, we selected 3 representative results from all 6 parameter sets we used in FPM and trial 1 is the best result we got. We note that when ground truth is unknown, all FPM reconstruction results might be treated as the "correct" solution as they all possess good contrast and fine details. However, apparent discrepancies are noticeable when comparing these results with the ground truth image. APIC, as an analytical method, is not prone to parameter selection.

As shown by the result, the reconstructed finer spokes were distorted in all reconstruction results of FPM. Moreover, noticeable wavy reconstruction artefacts existed in the phases reconstructed by FPM. When the measurements were given to APIC, the reconstructed phases and amplitudes were less wavy. The reconstructed amplitude is also closer to the ground truth, which is sampled using a high-NA objective. We stress that when the ground truth is not given, all the three FPM results shown in Fig. 2.6 may be perceived as a "good" reconstruction in practice as they preserve good contrast and are detail rich. However, these reconstructions are of low fidelity as they all deviate from the ground truth we acquired. We can find some of the spokes of the Siemens star target were missing in the FPM's reconstruction, which indicates the failure of FPM. This experiment showcased the ability of APIC to better retrieve a high-resolution complex field when the raw data size is constrained because it is an analytical method and does not rely as heavily on pupil overlap redundancy for solution convergence that FPM requires.

It is also worth noting, for input image tile of length 256 pixels on both sides, APIC reconstruction took 9 seconds on a personal computer (CPU: Intel Core i9-10940X with 64 GB RAM), while FPM required 25 seconds to finish the reconstruction. The relative computational efficiency of APIC can again be attributed to the analytical nature of its approach in contrast to FPM. We note that this computational efficiency is image tile size dependent — the smaller the tile is, the more efficient APIC can be (see Section 5 in the supplementary note for more information). As it is generally preferred to divide a large image into more smaller tiles in parallel computing, APIC's computational efficiency for smaller tiles aligns well with practical computation considerations.

In the next experiment, we studied the robustness of APIC and FPM at addressing optical aberrations. For this experiment, we acquired in total 316 images, which consisted of 52 normal brightfield measurements, 16 NA-matching measurements and 248 darkfield measurements. The nominal scanning pupil overlap ratio of our dataset was approximately 87% and the final theoretical synthetic NA was equal to 0.75 when all darkfield measurements were used. We note that this large degree of spectrum overlap was chosen to provide sufficient data redundancy for the best performance of FPM. APIC does not require such a large dataset (examples can be found in Fig. 2.6, and Fig. 2.14 in Section 2.6). In our reconstruction, APIC only used the NA-matching and darkfield measurements, whereas FPM used the entire dataset, including these additional 52 brightfield measurements corresponding to illumination angles that were below the objective's acceptance angle.

We deliberately defocused a Siemens star target to assess how the two methods perform under different aberration levels. In this experiment, the sample was defocused to different levels and the defocus information was hidden from both methods. The results of FPM and APIC are shown in Fig. 2.7a. Clearly, for large aberrations whose phase standard deviation exceeded 1.1π (the case when Siemens star target was defocused by 32 µm, and the maximal phase difference is approximately 3.8π), FPM failed to find the correct solution and the reconstructed images were considerably different from the ground truth, even when the algorithm indicated its convergence criterion was reached. At a lower aberration level, the amplitude reconstructions of FPM appeared to be close to the ideal case. However, the reconstructed phases were substantially different from the result when no defocus was introduced. In contrast, APIC was highly robust to different levels of aberrations. Although the contrast of APIC's reconstruction dropped under larger aberrations, it retrieved the correct aberrations and gave high-resolution complex field reconstructions that match with the in-focus result. The measured resolution for both FPM and APIC is approximately 870 nm when the in-focus measurements were used, which is close to the 840 nm theoretical resolution (Fig. 2.15).

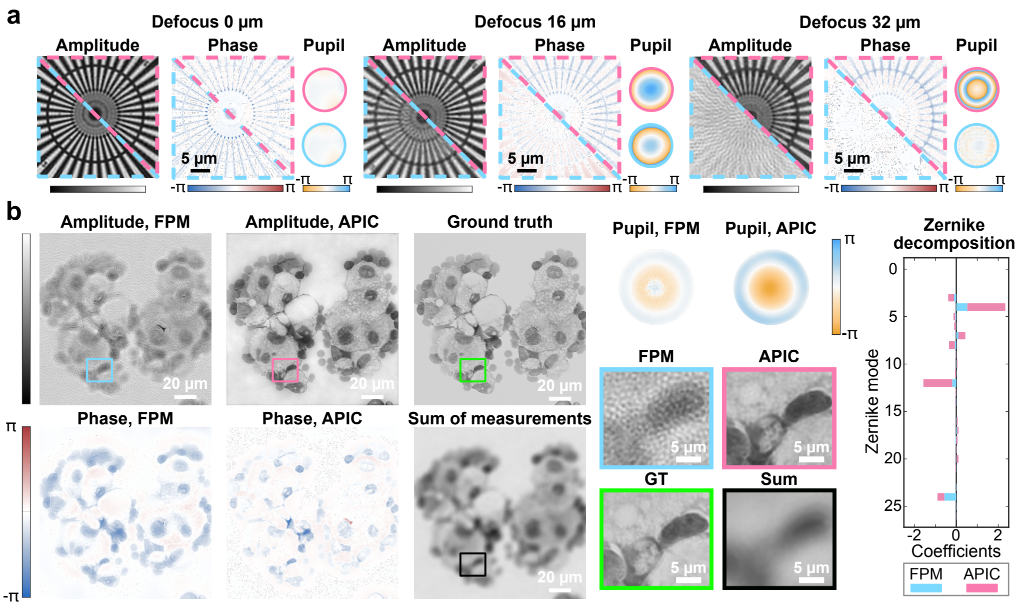

Figure 2.7: APIC under different levels of aberrations. **a**, Reconstructed complex fields and aberrations with different defocus distances. For the reconstruction of APIC and FPM, the defocus distance is labeled on top of each group. In our reconstruction, the actual defocus distance is hidden from both algorithms. APIC reconstructed amplitude and phase are shown on the upper right of each group and highlighted by the dashed magenta line. FPM reconstructed amplitude and phase are shown on the lower left of each group and highlighted by the dash cyan line. **b**, Reconstruction of a human thyroid carcinoma cell sample using APIC and FPM. Sum of measurements denotes the summation of all 316 images we acquired, which can be treated as the incoherent image we would get under the same objective. The ground truth was acquired using an objective whose magnification power is 40 and NA equals 0.75. The NA of this 40x objective equals the theoretical synthetic NA of APIC and FPM. We calculated the square root of the summed image and the 40x image to match up with the amplitude reconstruction. The zoomed images of the highlighted boxes are shown on the lower right of **b**. The Zernike decompositions of retrieved aberrations using FPM and APIC are shown on the far-right side of **b**.

To test the two methods under more complex aberration, we used an obsolete Olympus objective (10x magnification, NA 0.25) that was designed to work with another type of tube lens for image measurement in this particular experiment. A human thyroid adenocarcinoma cell sample was imaged to see their performances. As the standard deviation of the phase of the imaging system's aberration was close to $2\pi/5$, FPM failed to reconstruct a high-resolution image. From Fig. 2.7b, the reconstructed amplitude of FPM was heavily distorted by the reconstruction artefacts. APIC recovered all the finer details that were in good correspondence with the image we acquired using a 0.75 NA objective.

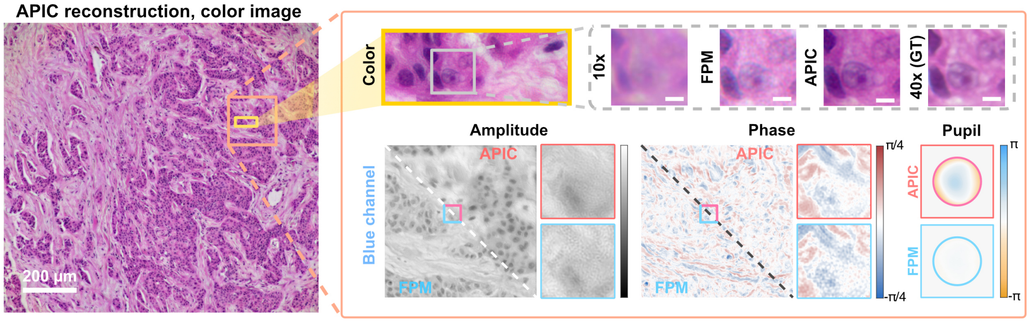

Figure 2.8: APIC result of hematoxylin and eosin (H&E) stained breast cancer cells. APIC reconstructed color image is shown on the left. The zoomed image of the highlighted region in the color image is shown on the right. The image with label "10x" denotes the image acquired using the same 10x magnification objective which was used for data acquisition. The color image with label "40x (GT)" denotes the ground truth we acquired using a 40x objective whose NA equals the theoretical synthetic NA of APIC. We note that we manually focused the image under red, green, and blue LED illumination when acquiring the ground truth as the best focal planes for them are different. We picked blue channel of an example in this illustration and the complex field reconstructions and retrieved aberrations are shown on the bottom of the rounded box. Scale bar for the zoomed color images: 5 μm.

We then conducted an experiment using a hematoxylin and eosin (H&E) stained breast cancer cell sample. We used red, green and blue LEDs to acquire dataset for these three different channels and then applied APIC for the reconstruction. In this experiment, the sample was placed at a fixed height in the data acquisition process. As a result, we see different levels of defocus in different channels lying on top of the chromatic aberrations of the objective (Fig. 2.16). To acquire the ground truth image, we switched to a 40x objective and manually focused each channel. We calibrated the illumination angles for the central patch (side length: 512 pixels) and then calculated the angles for off-axis patches using geometry. This calibrated illumination angles were used as input parameter data in our reconstruction.

The reconstructed color image is shown in Fig. 2.8. The comparison of all three channels can be found in Section 2.6 (Fig. 2.16). From the zoomed images in Fig. 2.8, we can see that the reconstructions of FPM were noisy for the blue channel. We found that FPM did not work well with this weakly absorptive sample under a relatively high aberration level. It failed to extract the aberration of the imaging system. As such, the color image generated by FPM appeared grainy and the high spatial frequency information were only partially recovered. We also see that the color reconstruction of APIC retained all high spatial frequency features that were closely matched up with the ground truth we acquired. This demonstrates that the aberration and complex field reconstruction of APIC is considerably more accurate compared with FPM.

2.4 The strengths of APIC and discussion on possible improvements

We showed that APIC can extract large aberrations and synthesize a large FOV and high-resolution images using low NA objectives. APIC empowers computational label-free microscopy with high robustness against aberration. Under the same high-aberration conditions, FPM fails to recover the aberration and its reconstruction result largely inherits such aberration and thus cannot produce aberration-free, high-resolution reconstructions.

Moreover, some of the fundamental problems in the conventional phase retrieval algorithm, such as being prone to optimization parameters and getting stuck in local minimum, are solved in APIC. Previous results demonstrated that reconstruction artifacts appear in FPM without properly selected parameters or loss functions[2, 9], which is in consistent with our experiment results shown in Fig. 2.6. Without a properly engineered metric, the selection of parameters becomes highly subjective. This again indicates that it is often unclear on whether FPM is even converging close to the real complex field solution. APIC is robust against this problem as it does not require an iterative algorithm for reconstruction. It circumvents the need of choosing optimization parameters or designing metric for convergence. However, as an analytical method, the knowledge about position of the LEDs, as well as the alignment of the NA-matching angle illuminations, is important in APIC. When large calibration errors show up, the solution of APIC will be negatively impacted (Fig. 2.25). Thus, good calibration is required in APIC to get the correct solution.

As APIC directly solves for the complex field, it avoids the potentially timeconsuming iterative process. When a reasonable image patch size is chosen, APIC is computationally more efficient compared with FPM. As such, APIC alleviates the lengthy processing in FPM, making it a more appealing method (Fig. 2.17 in Section 2.6).

While this work demonstrates a working APIC prototype, we note that a key aspect of the prototype would need further design improvements if a larger field of view is desired. Specifically, in this prototype, we treat the LED illumination as a plane wave at the sample plane. However, for large field of views, the angle can be quite different for different patches in the entire FOV. This may lead to noisy reconstruction of NA-matching measurements, as previous study has indicated. We anticipate that this problem can be mitigated by increasing the distance of the LED and the sample. It can also be solved by designing better LED illumination systems. Additionally, the number of LED illuminations can be reduced in future systems by decreasing the overlap of two measured spectrums.

In conclusion, we demonstrate that APIC can provide high-resolution and large field-of-view label-free imaging with unprecedent robustness to aberrations. As an analytical method, APIC is insensitive to parameter selections and can compute the correct imaging field without getting trapped in local minimums. APIC's analyticity is particularly important in a range of exacting applications, such as digital pathology, where even minor errors are not tolerable. APIC guarantees the correct solution while FPM like iterative methods cannot. Additionally, APIC brings new possibilities to label-free computational microscopy as it affords greater freedom in the use of engineered pupils for various imaging purposes. We anticipate the APIC concept be can fruitfully adopted for other methods, such as the aberration found by APIC can potentially be used to correct incoherent imaging. The idea of using the known spectrum to reconstruct unknown spectrum can be readily adapted for use in other scenarios.

2.5 Theoretical framework of APIC

In this section, we present the mathematical form of the forward model and reconstruction step of APIC. We begin by defining the forward model and introducing the notations in our derivation. Subsequently, we present the mathematical derivation of APIC, which involves three main steps: reconstructing the field using NA-matching measurements (measurements acquired under tilted illumination whose angle matches with the maximal acceptance angle of an imaging system), extracting and correcting aberrations, and reconstructing the complex field using darkfield measurements.

In our derivation, we first present the Kramers-Kronig method, which enables the reconstruction of complex fields from NA-matching measurements [17–20]. Using these reconstructed fields, we analytically retrieve the aberration of the imaging system. The extracted aberration is then applied to correct the currently reconstructed aberrated fields and the subsequent darkfield associated field reconstructions. The aberration-corrected reconstructed complex field serves as our initial *a priori* knowledge, referred to as the known field, and its Fourier transform as the known spectrum. To achieve high-resolution imaging, we incorporate darkfield measurements and use them to expand the reconstructed sample spectrum in an orderly way. At each sub-step, we reconstruct the spectrum corresponding to an unused darkfield measurement whose illumination angle is the smallest among all remaining unused measurements. This newly reconstructed spectrum, together with the original reconstructed spectrum, serves as the new *a priori* knowledge in subsequent reconstructions. To recover the field associated with one darkfield measurement, we focus on the spatial frequency space and form a linear equation with respect to the unknown spectrum. By solving this equation, we obtain a closed-form solution of the unknown spectrum sampled in this darkfield measurement. Adding this newly reconstructed spectrum effectively expands the sample spectrum coverage to achieve higher resolution.

After all measurements are reconstructed, we obtain a high-resolution, aberration-free complex field reconstruction in a purely analytical manner.

The forward model

When a thin sample is illuminated by a plane wave emitted by the ith ($i = 1, 2, \cdots, n$) LED with a transverse k-vector $\mathbf{k}_i$ and then imaged by an optical system, the modulated sample spectrum is given by

$$\hat{S}_i(\mathbf{k}) = \hat{O}(\mathbf{k} - \mathbf{k}_i)H(\mathbf{k}), \tag{2.1}$$

where H is the coherent transfer function (CTF) associated with the imaging system, $\hat{O}$ is the original sample's spectrum, $\hat{S}_i$ is the ith measured spectrum, and $\mathbf{k}$ is the 2D (transverse) spatial frequency vector. Reciprocally, we denote the spatial coordinate on the sample plane with $\mathbf{x}$. Without loss of generality, we assume $\hat{O}(\mathbf{0})$ is a real number, as the absolute average phase delay is not our interest. For a thin sample, we can further assume that the majority of the incident illumination light does not change its original direction. That is, at any position **x** we have

$$\left|\left(\mathcal{F}^{-1}\left[\hat{O}(\mathbf{k})\delta(\mathbf{k})\right]\right)(\mathbf{x})\right| > \left|\left[\mathcal{F}^{-1}\left(\hat{O}(\mathbf{k})\left[1-\delta(\mathbf{k})\right]\right)\right](\mathbf{x})\right|, \tag{2.2}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform, $\left[\mathcal{F}^{-1}(\cdot)\right](\mathbf{x})$ means the inverse Fourier transform is evaluated at $\mathbf{x}$, $|\cdot|$ gives the modulus of a complex number, and $\delta(\mathbf{k})$ is the "Kronecker" delta (We note that this is not a proper definition of a Kronecker delta function which is defined over integers. To be rigorous, one needs to define a function equals one in an ε-neighbor around the origin and zero otherwise. Then, we can proceed from that).

$$\delta(\mathbf{k}) = \begin{cases} 1, & \mathbf{k} = \mathbf{0}, \\ 0, & \text{otherwise}. \end{cases} \tag{2.3}$$

The physical meaning of this assumption is that the ballistic light exiting from the sample plane is dominant over the scattering light almost everywhere within the field of view.

As only the intensity of the light field is directly measured, the signal we get from a camera is

$$I_i(\mathbf{x}) = \left|\left[\mathcal{F}^{-1}(\hat{S}_i)\right](\mathbf{x})\right|^2 = \left|S_i(\mathbf{x})\right|^2, \tag{2.4}$$

where I_i is the ith image captured when lighting up the ith LED, and S_i denotes the ith sampled field in real space (the inverse Fourier transform of $\hat{S}_i$).

It is worth noting that this intensity measurement is insensitive to phase shift applied to the inverse Fourier transform

$$\left|\left[\mathcal{F}^{-1}(\hat{S}_i)\right](\mathbf{x})e^{j\xi(\mathbf{x})}\right|^2 = \left|\left[\mathcal{F}^{-1}(\hat{S}_i)\right](\mathbf{x})\right|^2 = I_i(\mathbf{x}), \tag{2.5}$$

where $\xi(\mathbf{x})$ stands for an arbitrary phase function and j is the unit imaginary number. Thus, we can choose a particular phase ramp, namely $-2\pi\mathbf{k}_i \cdot \mathbf{x}$, which effectively shifts $\hat{S}_i$ along the opposite direction of $\mathbf{k}_i$. Using the properties of Fourier transform when applying this special phase ramp, the measured intensity image is identical to the intensity of the inverse Fourier transform of the following (translated) spectrum

$$\hat{S}_i'(\mathbf{k}) = \left[\mathcal{F}(S_i')\right](\mathbf{k}) = \left[\hat{S}_i * \mathcal{F}(e^{-j2\pi\mathbf{k}_i\cdot\mathbf{x}})\right](\mathbf{k}) = \hat{O}(\mathbf{k})H(\mathbf{k}+\mathbf{k}_i), \tag{2.6}$$

where $\cdot$ is the dot product, $S_i'(\mathbf{x}) := S_i(\mathbf{x})e^{-2\pi j\mathbf{k}_i\cdot\mathbf{x}}$, and $*$ is convolution. This model suggests that different components of sample's spectrum can be measured with different illumination angles. By increasing the illumination angle, we can effectively sample high spatial frequency information which is normally inaccessible due to limited numeric aperture (NA) of an imaging system.

The CTF support of an imaging system can be reasonably assumed as circular with a NA-dependent radius k_{NA}, and its phase ϕ of the CTF function fully depicts the system's aberrations. For simplicity, we abuse the notation and use NA to represent this radius k_{NA} in the following section. We then have

$$\begin{aligned} H(\mathbf{k}) &= \mathrm{Circ}_{\mathrm{NA}}(\mathbf{k})e^{j\phi(\mathbf{k})} = \mathbb{1}(|\mathbf{k}| \leq \mathrm{NA})e^{j\phi(\mathbf{k})} \\ &= \begin{cases} e^{j\phi(\mathbf{k})}, & \text{if } |\mathbf{k}| \leq \mathrm{NA}, \\ 0, & \text{otherwise,} \end{cases} \end{aligned} \tag{2.7}$$

where $\phi(\mathbf{k})$ stands for the system's aberration, and $\mathrm{Circ}_{\mathrm{NA}}(\mathbf{k}) := \mathbb{1}(|\mathbf{k}| \leq \mathrm{NA})$ is an indicator function that gives one when the modulus of the k-vector is below the system NA and zero otherwise.

We further decompose $\hat{O}(\mathbf{k})$ to its amplitude and phase. Together with Eq. 2.7, we can rewrite Eq. 2.6, which gives

$$\hat{S}_i'(\mathbf{k}) = \hat{O}(\mathbf{k})H(\mathbf{k}+\mathbf{k}_i) = \hat{A}(\mathbf{k})e^{j\hat{\alpha}(\mathbf{k})}\,\mathrm{Circ}_{\mathrm{NA}}(\mathbf{k}+\mathbf{k}_i)e^{j\phi(\mathbf{k}+\mathbf{k}_i)}, \tag{2.8}$$

where $\hat{A}(\mathbf{k}) := \left|\hat{O}(\mathbf{k})\right| \in \mathbb{R}$ is the amplitude of the sample's spectrum and $\hat{\alpha}(\mathbf{k}) := \arg\left[\hat{O}(\mathbf{k})\right] \in \mathbb{R}$ is its phase (the operator arg gives the argument of a complex number).

In the following two subsections, we proceed with Eq. 2.6 and show that the closed-form solution of the sample's spectrum $\hat{O}(\mathbf{k})$ can be obtained (Sections 2.5 and 2.5), and the aberration of the imaging system $\phi(\mathbf{k})$ can also be analytically retrieved (Section 2.5). This is achieved by only using the NA-matching and darkfield measurements. For simplicity, in the following discussion, we assume all $|\mathbf{k}_i| \geq \mathrm{NA}$ and $\mathbf{k}_i$ is ordered so that $|\mathbf{k}_i| \leq |\mathbf{k}_{i+1}|$.

Reconstruction under the NA-matching angle illumination

For integrity, we reproduce the main result of recent works on spatial Kramers-Kronig relations which enables complex field reconstruction when the illumination angle exactly matches with the maximal acceptance angle of the objective [17–19]. However, the aberration consideration is missing in these existing works. Here, we take the imaging system's aberration into consideration. The main result can be found in Eqs. 2.32 and 2.33. In this subsection, we only consider the first n_0 measurements whose k-vector $\mathbf{k}_i$ satisfies

$$|\mathbf{k}_i| = \mathrm{NA}, \quad i = 1, 2, \cdots, n_0, \quad \text{where } n_0 := \max\left\{i \in \mathbb{N} \,\middle|\, |\mathbf{k}_i| = \mathrm{NA}\right\}. \tag{2.9}$$

We call these measurements as the NA-matching measurements. Instead of presenting this result using directional Hilbert transform and 1D Kramers-Kronig relations, we take the opportunity to put more focus on the underlying structure of the (logarithm mapped) signal's Fourier transform. To be specific, we will construct a signal whose Fourier transform is one-sided (which will be defined later), and show that this allows us to calculate the imaginary part of the signal from its real part based on the fact that Fourier transform is a linear operator. This carefully constructed signal can be mapped back to the complex field we want. We note that the approach presented here can also be generalized to higher dimensional spaces in an intuitive way. Although this method allows analytical complex field reconstruction, we found that both our desired sample's field and the aberration function of the imaging system are entangled in the reconstructed field.

As our goal is computing the imaginary part of a complex signal from its real part using Kramers-Kronig relations, the real part must be known first. However, we only measure the intensity (squared modulus) of a complex field, which is neither its real or imaginary part. To solve this mismatch, we apply a nonlinear transformation that maps the intensity and phase of a complex number to the real and imaginary part of its output, respectively. This can be done by taking the logarithm of a nonzero complex number. Applying this to the complex field $S_i'(\mathbf{x})$ in a point-wise manner, we have

$$\log\left[S_i'(\mathbf{x})\right] = \log\left[\left|S_i'(\mathbf{x})\right|\right] + j\arg\left[S_i'(\mathbf{x})\right]. \tag{2.10}$$

The nonzero condition is guaranteed by Eq. 2.2. Note that the first term on the right-hand side is purely real and the second term is purely imaginary. As the intensity $I_i(\mathbf{x}) = \left|S_i(\mathbf{x})\right|^2 = \left|S_i(\mathbf{x}) \cdot e^{-2\pi j \mathbf{k}_i \cdot \mathbf{x}}\right|^2 = \left|S_i'(\mathbf{x})\right|^2$ is measured, we know everything about the real part of $\log\left[S_i'(\mathbf{x})\right]$. What remains is to verify that its imaginary part can be reconstructed using this known real part.

In the remaining part of this section, we will show that the Fourier transform of this transformed signal is one-sided. Such structure allows us to perform our desired reconstruction. We first give the definition of a signal being one sided.

Definition 2.1 *For any positive integer* $m \in \mathbb{Z}^+$, *we say* $g : \mathbb{R}^m \rightarrow \mathbb{C}$ *is one sided if there exists a nonzero vector* $\mathbf{e} \in \mathbb{R}^m, \mathbf{e} \neq \mathbf{0}$ *such that its Fourier transform* $\hat{g}$ *satisfies*

$$\hat{g}(\mathbf{k}) = 0, \quad \forall \mathbf{k} \in \mathbb{R}^m \ s.t.\ \mathbf{k} \cdot \mathbf{e} < 0. \tag{2.11}$$

We say g *is strictly one sided if* $\hat{g}(\mathbf{k}) = 0,\ \forall \mathbf{k} \in \mathbb{R}^m\ s.t.\ \mathbf{k} \cdot \mathbf{e} \leq 0.$

To show the transform signal is one sided, we are going to factor out the "offset" field, $\mathcal{R}(\mathbf{x})$, which is the field that does not change its direction

$$\mathcal{R}(\mathbf{x}) = \left(\mathcal{F}^{-1}\left[\hat{S}_i'(\mathbf{k})\delta(\mathbf{k})\right]\right)(\mathbf{x}) = \left(\mathcal{F}^{-1}\left[\hat{O}(\mathbf{k})H(\mathbf{k}+\mathbf{k}_i)\delta(\mathbf{k})\right]\right)(\mathbf{x}) = re^{j\phi(\mathbf{k}_i)}, \tag{2.12}$$

where r is real. When we factor out this field from $S_i'(\mathbf{x})$, we have

$$\frac{S_i'(\mathbf{x})}{\mathcal{R}(\mathbf{x})} = \frac{S_i'(\mathbf{x})}{re^{j\phi(\mathbf{k}_i)}} = \frac{S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}}{r}. \tag{2.13}$$

We can see that $\phi(\mathbf{k}_i)$ serves as a phase offset and r can be treated as a normalization factor. Note that r and $\phi(\mathbf{k}_i)$ are unknown in reality, and only the intensity of $S_i(\mathbf{x})$ is measured. Thus, we primarily focus on $S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}$ in our following derivation because adding a phase offset to a complex field does not change its intensity.

Applying the logarithm transformation to the offset field $S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}$, we have

$$\begin{aligned}\log\left[S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}\right] &= \log\left[\frac{S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}}{r}\right] + \log(r) \\ &= \log\left[1 + \frac{S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)} - r}{r}\right] + \log(r).\end{aligned} \tag{2.14}$$

With the assumption that the majority of the light does not change its direction (Eq. 2.2), we have the offset field is larger than the sample modulated field $r = \left|\mathcal{R}(\mathbf{x})\right| > \left|S_i'(\mathbf{x}) - \mathcal{R}(\mathbf{x})\right| = \left|S_i'(\mathbf{x}) - re^{j\phi(\mathbf{k}_i)}\right|$. So, we have $r > \left|S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)} - r\right|\left|e^{j\phi(\mathbf{k}_i)}\right| = \left|S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)} - r\right|$. Thus, we can write the first term on the right-hand side in a convergent Taylor series (we will see this makes it easier to analyze the structure of its Fourier transform)

$$T(\mathbf{x}) := \log\left[1 + \frac{S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)} - r}{r}\right] = \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \frac{\left[S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)} - r\right]^m}{r^m}. \tag{2.15}$$

Note this applies to all $\mathbf{x}$. For simplicity, we define $\Delta_i(\mathbf{x})$ as

$$\begin{aligned}\Delta_i(\mathbf{x}) &:= S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)} - r = \left[S_i'(\mathbf{x}) - re^{j\phi(\mathbf{k}_i)}\right]e^{-j\phi(\mathbf{k}_i)} \\ &= \left[S_i'(\mathbf{x}) - \mathcal{R}(\mathbf{x})\right]e^{-j\phi(\mathbf{k}_i)},\end{aligned} \tag{2.16}$$

and Eq. 2.14 can be rewritten as

$$\log\left[S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}\right] = \log(r) + T(\mathbf{x}) = \log(r) + \sum_{m=1}^{\infty} \frac{(-1)^{m+1}}{m} \frac{\Delta_i^m(\mathbf{x})}{r^m}. \tag{2.17}$$

The Fourier transform of $\Delta_i(\mathbf{x})$ yields

$$\begin{aligned}\hat{\Delta}_i(\mathbf{k}) &= \left[\mathcal{F}(\Delta_i)\right](\mathbf{k}) = \left[\mathcal{F}\left(\left[S_i'(\mathbf{x}) - \mathcal{R}(\mathbf{x})\right]e^{-j\phi(\mathbf{k}_i)}\right)\right](\mathbf{k}) \\ &= \hat{S}_i'(\mathbf{k})\left[1 - \delta(\mathbf{k})\right]e^{-j\phi(\mathbf{k}_i)}.\end{aligned} \tag{2.18}$$

Lemma 2.1 *For* $\mathbf{x} \in \mathbb{R}^2$*, let* $f_1(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{C}$ *and* $f_2(\mathbf{x}) : \mathbb{R}^2 \to \mathbb{C}$ *be two complex* l_2 *function with Fourier transform* $\hat{f}_1(\mathbf{k})$ *and* $\hat{f}_2(\mathbf{k})$*, respectively. Assume there exists a (common) nonzero vector* $\mathbf{e} \in \mathbb{R}^2$*,* $\mathbf{e} \neq \mathbf{0}$ *such that*

$$\hat{f}_1(\mathbf{k}) = \hat{f}_2(\mathbf{k}) = 0, \forall \mathbf{k} \text{ } s.t. \text{ } \mathbf{k} \cdot \mathbf{e} < 0, \tag{2.19}$$

then their product $f'(\mathbf{x}) = f_1(\mathbf{x}) f_2(\mathbf{x})$ *is one sided. Furthermore, if* f_1 *and* f_2 *are strictly one sided (with the same* $\mathbf{e}$*), their product is also strictly one sided.*

Proof: We take the Fourier transform on both sides, which yields

$$\hat{f}'(\mathbf{k}) = \left[\mathcal{F}(f')\right](\mathbf{k}) = \left[\mathcal{F}(f_1 f_2)\right](\mathbf{k}) = \Big(\left[\mathcal{F}(f_1)\right] * \left[\mathcal{F}(f_1)\right]\Big)(\mathbf{k}) = \left[\hat{f}_1 * \hat{f}_2\right](\mathbf{k}). \tag{2.20}$$

For $\mathbf{k}_0 \in \mathbb{R}^2$ such that $\mathbf{k}_0 \cdot \mathbf{e} < 0$ ($\mathbf{e} \neq \mathbf{0}$), we have

$$\begin{aligned} \hat{f}'(\mathbf{k}_0) = \left[\hat{f}_1 * \hat{f}_2\right](\mathbf{k}_0) &= \int d\mathbf{k}' \hat{f}_1(\mathbf{k}') \hat{f}_2(\mathbf{k}_0 - \mathbf{k}') \\ &= \int_{\mathbf{k}' \cdot \mathbf{e} < 0} d\mathbf{k}' \hat{f}_1(\mathbf{k}') \hat{f}_2(\mathbf{k}_0 - \mathbf{k}') + \int_{\mathbf{k}' \cdot \mathbf{e} \geq 0} d\mathbf{k}' \hat{f}_1(\mathbf{k}') \hat{f}_2(\mathbf{k}_0 - \mathbf{k}'). \end{aligned} \tag{2.21}$$

If $\mathbf{k}' \cdot \mathbf{e} < 0$, we have $\hat{f}_1(\mathbf{k}') = 0$. If $\mathbf{k}' \cdot \mathbf{e} \geq 0$, we have $(\mathbf{k}_0 - \mathbf{k}') \cdot \mathbf{e} = \mathbf{k}_0 \cdot \mathbf{e} - \mathbf{k}' \cdot \mathbf{e} < 0$. As f_1 and f_2 are both one sided and share the same $\mathbf{e}$, we can conclude

$$\begin{aligned} \hat{f}'(\mathbf{k}_0) &= \int_{\mathbf{k}' \cdot \mathbf{e} < 0} d\mathbf{k}' \hat{f}_1(\mathbf{k}') \hat{f}_2(\mathbf{k}_0 - \mathbf{k}') + \int_{\mathbf{k}' \cdot \mathbf{e} \geq 0} d\mathbf{k}' \hat{f}_1(\mathbf{k}') \hat{f}_2(\mathbf{k}_0 - \mathbf{k}') \\ &= 0, \quad \forall \mathbf{k}_0 \text{ s.t. } \mathbf{k}_0 \cdot \mathbf{e} < 0. \end{aligned} \tag{2.22}$$

That is, f' is one sided. The proof for f_1 and f_2 being strictly one sided follows the same structure. To see that, we calculate the integration for $\mathbf{k}' \cdot \mathbf{e} \leq 0$ and $\mathbf{k}' \cdot \mathbf{e} > 0$. Then, we can prove the strict version with the same technique.

We can easily see that $\Delta_i(\mathbf{x})$ is one sided. To show this, we choose $\mathbf{e} = -\mathbf{k}_i$. As the illumination angle is matched with NA (Eq. 2.9), we can verify that for arbitrary $\mathbf{k}$ such that $\mathbf{k} \cdot \mathbf{e} < 0$, we have

$$|\mathbf{k} + \mathbf{k}_i| = \sqrt{|\mathbf{k}|^2 + |\mathbf{k}_i|^2 + 2\mathbf{k} \cdot \mathbf{k}_i} = \sqrt{|\mathbf{k}|^2 + |\mathbf{k}_i|^2 - 2\mathbf{k} \cdot \mathbf{e}} > \sqrt{|\mathbf{k}_i|^2} = \text{NA}. \tag{2.23}$$

Note the last equality holds because we are interested in the NA-matching angle illumination condition (Eq. 2.9) in this section (Section 2.5). If we evaluate the

CTF function H at $\mathbf{k}+\mathbf{k}_i$, we have $H(\mathbf{k}+\mathbf{k}_i) = \mathbb{1}(|\mathbf{k}+\mathbf{k}_i| \leq \mathrm{NA})e^{j\phi(\mathbf{k})} = 0$ for all $\mathbf{k}$ subject to $\mathbf{k}_0 \cdot \mathbf{e} = \mathbf{k} \cdot (-\mathbf{k}_i) < 0$. We can then prove Δ_i is one sided:

$$\begin{aligned}\hat{\Delta}_i(\mathbf{k}) &= \hat{S}_i'(\mathbf{k})\big[1-\delta(\mathbf{k})\big]e^{-j\phi(\mathbf{k}_i)} = \hat{O}_i(\mathbf{k})H(\mathbf{k}+\mathbf{k}_i)\big[1-\delta(\mathbf{k})\big]e^{-j\phi(\mathbf{k}_i)} \\ &= 0, \qquad \forall \mathbf{k} \text{ s.t. } \mathbf{k}\cdot(-\mathbf{k}_i) < 0. \end{aligned} \tag{2.24}$$

Furthermore, for $\mathbf{k}\cdot\mathbf{e} = -\mathbf{k}\cdot\mathbf{k}_i = 0$, we have $\mathbf{k} \perp \mathbf{k}_i$. When $\mathbf{k} \neq \mathbf{0}$, the above two equations still hold. We only need to consider the special case, namely $\mathbf{k} = 0$. At $\mathbf{k} = 0$, we have

$$\hat{\Delta}_i(\mathbf{0}) = \hat{S}_i'(\mathbf{0})\big[1-\delta(\mathbf{0})\big]e^{-j\phi(\mathbf{k}_i)} = \hat{S}_i'(\mathbf{0})(1-1)e^{-j\phi(\mathbf{k}_i)} = 0. \tag{2.25}$$

That is, Δ_i is strictly one sided. Using lemma 2.1, we conclude the Taylor series (the last term on the right-hand side of Eq. 2.17) is strictly one sided by induction.

Note that if one wants to directly use S_i instead, the "offset" field should be modified as $re^{2\pi j\mathbf{k}_i\cdot\mathbf{x}+j\phi(\mathbf{k}_i)}$. This is because the original model uses tilted illumination, so the unchanged light field is associated with an incident angle. With such modification, one can verify it gives an identical result.

Here, we have proved that the transformed signal $\log\big[S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}\big]$ is one sided (It is easy to see that $\log\big[S_i'(\mathbf{x})\big]$ is also one sided, as $e^{-j\phi(\mathbf{k}_i)}$ is a constant). We proceed to show that such special structure allows one to reconstruct the imaginary part from the real part of the signal.

We now get more involved with a modified version of Eq. 2.10

$$\begin{aligned}\log\big[S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}\big] &= \log\big[\big|S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}\big|\big] + j\arg\big[S_i'(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}\big] \\ &= \log\big[\big|S_i'(\mathbf{x})\big|\big] \qquad + j\Big(\arg\big[S_i'(\mathbf{x})\big] - \phi(\mathbf{k}_i)\Big) \\ &= \log(r) + \Re\big[T(\mathbf{x})\big] \quad + j\Im\big[T(\mathbf{x})\big], \end{aligned} \tag{2.26}$$

where the last equality follows from Eq. 2.17, $\Re(\cdot)$ denotes the real part of a complex number, and $\Im(\cdot)$ denotes the imaginary part. The Fourier transform exhibits even symmetric for a real signal and odd symmetric for an imaginary signal, so we have

$$\big[\mathcal{F}\big(\log\big|S_i'\big|\big)\big](\mathbf{k}) = \big[\mathcal{F}\big(\log\big|S_i'\big|\big)\big]^*(-\mathbf{k}), \tag{2.27}$$

and

$$\Big[\mathcal{F}\Big(j\big[\arg(S_i') - \phi(\mathbf{k}_i)\big]\Big)\Big](\mathbf{k}) = -\Big[\mathcal{F}\Big(j\big[\arg(S_i') - \phi(\mathbf{k}_i)\big]\Big)\Big]^*(-\mathbf{k}), \tag{2.28}$$

where $*$ in the superscript denotes complex conjugate. Because we have already proved that $\log\left[S'_i(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}\right]$ is one sided, we have

$$\begin{aligned}\left[\mathcal{F}\Big(\log[S'_i e^{-j\phi(\mathbf{k}_i)}]\Big)\right](\mathbf{k}) &= \left[\mathcal{F}\big(\log\left|S'_i\right|\big)\right](\mathbf{k}) + \left[\mathcal{F}\Big(j\left[\arg(S'_i) - \phi(\mathbf{k}_i)\right]\Big)\right](\mathbf{k}) = 0\\ \Rightarrow \left[\mathcal{F}\Big(j\left[\arg(S'_i) - \phi(\mathbf{k}_i)\right]\Big)\right](\mathbf{k}) &= -\left[\mathcal{F}\big(\log\left|S'_i\right|\big)\right](\mathbf{k}), \quad \forall \mathbf{k} \text{ s.t. } \mathbf{k}\cdot(-\mathbf{k}_i) < 0.\end{aligned} \tag{2.29}$$

For $\mathbf{k}\cdot(-\mathbf{k}_i) > 0$, we can use symmetry to conclude

$$\left[\mathcal{F}\Big(j\left[\arg(S'_i) - \phi(\mathbf{k}_i)\right]\Big)\right](\mathbf{k}) = \left[\mathcal{F}\big(\log\left|S'_i\right|\big)\right](\mathbf{k}), \quad \forall \mathbf{k} \text{ s.t. } \mathbf{k}\cdot(-\mathbf{k}_i) > 0. \tag{2.30}$$

As $T(\mathbf{x})$ is strictly one sided, its Fourier transform is zero for all $\mathbf{k}$ such that $\mathbf{k} \perp \mathbf{k}_i$. Moreover, the Fourier transform of constant $\log(r)$ is a real (Dirac delta) function centered at zero. In other words, the Fourier transform of $\log(r)$ has no imaginary part. So, we have

$$\left[\mathcal{F}\Big(j\left[\arg(S'_i) - \phi(\mathbf{k}_i)\right]\Big)\right](\mathbf{k}) = 0, \quad \forall \mathbf{k} \text{ s.t. } \mathbf{k}\cdot(-\mathbf{k}_i) = 0. \tag{2.31}$$

By collecting all pieces from the above three equations and noticing $I_i(\mathbf{x}) = \left|S'_i(\mathbf{x})\right|^2$, we have

$$\begin{aligned}&\left[\mathcal{F}\Big(\log[S'_i e^{-j\phi(\mathbf{k}_i)}]\Big)\right](\mathbf{k})\\ &\qquad = \begin{cases} 2\left[\mathcal{F}\big(\log\left|S'_i\right|\big)\right](\mathbf{k}) = \left[\mathcal{F}\big(\log I_i\big)\right](\mathbf{k}), & \mathbf{k}\cdot\mathbf{k}_i < 0,\\ \left[\mathcal{F}\big(\log\left|S'_i\right|\big)\right](\mathbf{k}) = \frac{1}{2}\left[\mathcal{F}\big(\log I_i\big)\right](\mathbf{k}), & \mathbf{k}\cdot\mathbf{k}_i = 0,\\ 0, & \mathbf{k}\cdot\mathbf{k}_i > 0.\end{cases}\end{aligned} \tag{2.32}$$

That is, the complex field can be reconstructed using its real part $\log\left|S'_i\right|$. We can then restore the desired field (up to a constant phase offset) with inverse Fourier transform and apply exponential function to each point of the inverse Fourier transform

$$S'_i(\mathbf{x})e^{-j\phi(\mathbf{k}_i)} = \exp\left[\left(\mathcal{F}^{-1}\left[\mathcal{F}\Big(\log[S'_i e^{-j\phi(\mathbf{k}_i)}]\Big)\right]\right)(\mathbf{x})\right]. \tag{2.33}$$

Note that if there is no aberration in the system $\phi(\mathbf{k}) \equiv 0$, we have $S'_i(\mathbf{x}) = \hat{O}(\mathbf{k})\,\mathrm{Circ}_{\mathrm{NA}}(\mathbf{k}+\mathbf{k}_i)$ and the phase offset $\phi(\mathbf{k}_i) = 0$, which reduces to the result reported in previous literature [17–19].

Aberration extraction

In Eq. 2.33, we end up with the reconstruction of $S'_i(\mathbf{x})e^{-j\phi(\mathbf{k}_i)}$. Now we take a close look at its spectrum:

$$\begin{aligned}\hat{S}'_i(\mathbf{k})e^{-j\phi(\mathbf{k}_i)} &= \hat{O}(\mathbf{k})H(\mathbf{k}+\mathbf{k}_i)e^{-j\phi(\mathbf{k}_i)}\\ &= \hat{A}(\mathbf{k})e^{j\hat{\alpha}(\mathbf{k})}\,\mathrm{Circ}_{\mathrm{NA}}(\mathbf{k}+\mathbf{k}_i)e^{j\phi(\mathbf{k}+\mathbf{k}_i)-j\phi(\mathbf{k}_i)}.\end{aligned} \tag{2.34}$$

From this equation, we can see the sample's spectrum and the system's aberration function are superimposed in the reconstruction. For a practical optical system with aberration, we have not yet reconstructed a clean sample spectrum. Instead, we reconstructed an aberrated version of the desired spectrum.

Without correction, aberrations of an imaging system, including defocus due to sample's height unevenness, would largely degrade the reconstruction quality. Nonetheless, we will show that this problem can be tackled by working in the spatial frequency domain (that is, working with the spectrum). We prove that the aberration of the system is analytically solvable by considering the phases of the reconstructed spectrum.

To extract the aberration from this reconstructed spectrum, we will need a method to separate the contribution from the sample itself and the imaging system. We can achieve this by considering the phases of multiple reconstructed spectrums. We first define the overlap of two spectrums.

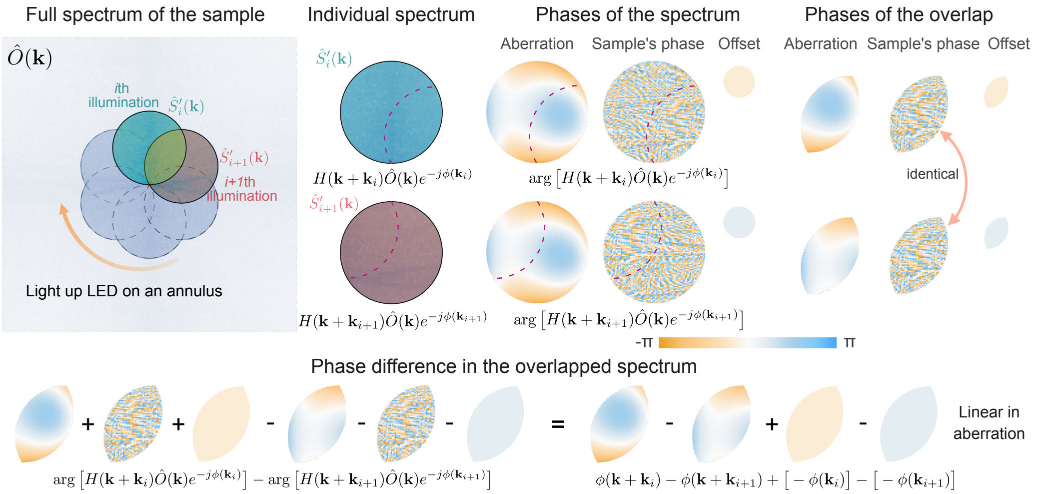

Figure 2.9: Phase difference of two spectrums with overlap. $\hat{O}(\mathbf{k})$ stands for the original sample's spectrum, $\hat{S}_i(\mathbf{k})$ is the acquired spectrum under ith illumination, and $H(\mathbf{k})$ is the coherent transfer function (CTF). For aberration extraction, we only use LED whose illumination k vector is matched with the maximal receiving angle of the imaging system $\mathbf{k}_i = \text{NA}$. Here we select $l = i + 1$ for illustration as the LEDs are lit up along an annulus shown in the figure. In the overlapped region, the phase of the object itself cancels out and phase difference of the shifted CTF solely contributes to the measured phase difference. This means the phase difference is linearly dependent on the aberration. Thus, we can design a operator that counts for the shift induced phase difference and use it to recover the imaging system's aberration.

Definition 2.2 *For two reconstructed spectrums* $\hat{S}_i'(\mathbf{k})$ *and* $\hat{S}_l'(\mathbf{k})$*, we define a set* C_{il} *as*

$$C_{il} := \left\{\mathbf{k} \in \mathbb{R}^2 \middle| H(\mathbf{k}+\mathbf{k}_i)H(\mathbf{k}+\mathbf{k}_l) \neq 0\right\}. \tag{2.35}$$

We say this two spectrums are overlapped if the set C_{il} *is nonempty* $C_{il} \neq \emptyset$*, and the overlap between* $\hat{S}_i'(\mathbf{k})$ *and* $\hat{S}_l'(\mathbf{k})$ *is* C_{il}*.*

Let us consider two spectrums $\hat{S}_i'(\mathbf{k})$ and $\hat{S}_l'(\mathbf{k})$ with (nonempty) overlap $C_{il} \neq \emptyset$. We calculate their phase difference within the overlapped region C_{il}. This gives

$$\begin{aligned}
&\arg\left[\hat{S}_i'(\mathbf{k})e^{-j\phi(\mathbf{k}_i)}\right] - \arg\left[\hat{S}_l'(\mathbf{k})e^{-j\phi(\mathbf{k}_l)}\right] \\
&\qquad = \left[\hat{\alpha}(\mathbf{k}) + \phi(\mathbf{k}+\mathbf{k}_i) - \phi(\mathbf{k}_i)\right] - \left[\hat{\alpha}(\mathbf{k}) + \phi(\mathbf{k}+\mathbf{k}_l) - \phi(\mathbf{k}_l)\right] \\
&\qquad = \phi(\mathbf{k}+\mathbf{k}_i) - \phi(\mathbf{k}_i) - \phi(\mathbf{k}+\mathbf{k}_l) + \phi(\mathbf{k}_l), \\
&\qquad = \left[\phi(\mathbf{k}+\mathbf{k}_i) - \phi(\mathbf{k}+\mathbf{k}_l)\right] - \left[\phi(\mathbf{k}_i) - \phi(\mathbf{k}_l)\right], \quad \text{if } \mathbf{k} \in C_{il}.
\end{aligned} \tag{2.36}$$

We will refer the first term $\phi(\mathbf{k}+\mathbf{k}_i)-\phi(\mathbf{k}+\mathbf{k}_l)$ on the right-hand side as the aberration difference, and the last term $\phi(\mathbf{k}_i) - \phi(\mathbf{k}_l)$ as the offset. When we consider the phase difference of the two spectrums, the contribution from the sample spectrum cancels out and the difference depends solely on the system's aberration. We can easily see that the remaining phase difference is linear with respect to the aberration function. As such, we can construct a linear operator that maps the aberration into phase differences of two overlapping spectrums. To do so, we first rearrange the 2D spectrum into a vector.

Definition 2.3 *Assume* $\mathbf{B}$ *is a* $m \times t$ *matrix, we define a "flattening" operator* $\mathrm{Flat}_{m,t}$ *which concatenates every column of* $\mathbf{B}$ *and produces a vector of length* mt*. If we use* $\mathcal{B}$ *to denote this flattened vector, we have*

$$\mathcal{B} = \mathrm{Flat}_{m,t}(\mathbf{B}) \in \mathbb{R}^{mt\times 1} \quad \textit{and} \quad \mathcal{B}[i_1 + (i_2-1)m] = \mathbf{B}(i_1, i_2). \tag{2.37}$$

We also define an inverse operator $\mathrm{Flat}_{m,t}^{-1}$*, which restores the matrix when applied to the flattened vector*

$$\mathbf{B} = \mathrm{Flat}_{m,t}^{-1}(\mathcal{B}). \tag{2.38}$$

Assume the sample spectrum we acquired is a $N \times N$ matrix, we write $\hat{\mathbf{S}}_i' \in \mathbb{C}^{N\times N}$ as the matrix form of the sampled spectrum, and $\boldsymbol{\phi} \in \mathbb{R}^{N\times N}$ as the matrix form of the aberration. The corresponding flattened version is then $\hat{\boldsymbol{\mathcal{S}}}_i' = \mathrm{Flat}_{N,N}(\hat{\mathbf{S}}_i') \in \mathbb{C}^{N^2\times 1}$, and $\boldsymbol{\varphi} = \mathrm{Flat}_{N,N}(\boldsymbol{\phi}) \in \mathbb{R}^{N^2\times 1}$, respectively.

We let $\mathbf{K}_u, \mathbf{K}_v \in \mathbb{R}^{N\times N}$ be the spatial frequency grid (with zero frequency at $[c_0, c_0] = \left(\lceil\frac{N}{2}\rceil, \lceil\frac{N}{2}\rceil\right)$, where $\lceil\frac{N}{2}\rceil$ is the smallest integer that is larger than $\frac{N}{2}$) for the sampled spectrum, where u and v denotes two orthogonal direction. Let $\boldsymbol{K}_u, \boldsymbol{K}_v \in \mathbb{R}^{N^2\times 1}$ be their flattened vectors $\boldsymbol{K}_u = \mathrm{Flat}_{N,N}(\mathbf{K}_u)$ and $\boldsymbol{K}_v = \mathrm{Flat}_{N,N}(\mathbf{K}_v)$. For simplicity, we let $\boldsymbol{K}(m), m \in \{1, 2, \cdots, N^2\}$ denotes k vector in grid, which is defined as

$$\boldsymbol{K}(m) := \big[\boldsymbol{K}_u(m), \boldsymbol{K}_v(m)\big]. \tag{2.39}$$

Additionally, we define an index set $\mathcal{I}_{il}$ by

$$\mathcal{I}_{il} := \Big\{m = 1, 2, \cdots, N^2 \Big| \boldsymbol{K}(m) \in C_{il}\Big\}. \tag{2.40}$$

We let $|\mathcal{I}_{il}|$ be the cardinal of set $\mathcal{I}_{il}$ and also abuse the notation such that $\mathcal{I}_{il}(m)$ indicate the mth smallest element in $\mathcal{I}_{il}$. For the transverse illumination k-vector $\mathbf{k}_i$ ($i = 1, 2, \cdots, n$), we denote its grid representation as $\boldsymbol{\kappa}_i$.

With this notation, we design a difference operator $\mathbf{D}_{il} \in \mathbb{R}^{|\mathcal{I}_{il}|\times N^2}$. When acting on $\boldsymbol{\varphi}$, we want this operator to calculate the aberration difference $\phi(\mathbf{k}+\mathbf{k}_i) - \phi(\mathbf{k}+\mathbf{k}_l)$ associated with $\hat{\mathcal{S}}'_i$ and $\hat{\mathcal{S}}'_l$ (each row corresponds to one specific choice of $\mathbf{k}$). Based on this principle, we assign value 1 to the place that corresponds to phase $\phi(\mathbf{k}+\mathbf{k}_i)$ and value -1 to the place that corresponds to phase $\phi(\mathbf{k}+\mathbf{k}_l)$ for every row of $\mathbf{D}_{il}$. That is to say, the operator $\mathbf{D}_{il}$ is constructed as

$$\mathbf{D}_{il}(i_1, i_2) = \begin{cases} 1, & \text{if } \boldsymbol{K}(i_2) = \boldsymbol{K}\big[\mathcal{I}_{il}(i_1)\big] + \boldsymbol{\kappa}_i \\ -1, & \text{if } \boldsymbol{K}(i_2) = \boldsymbol{K}\big[\mathcal{I}_{il}(i_1)\big] + \boldsymbol{\kappa}_l \\ 0, & \text{otherwise.} \end{cases} \tag{2.41}$$

We can see this operator indeed calculates the aberration difference

$$\mathbf{D}_{il}\boldsymbol{\varphi} = \begin{bmatrix} \phi\Big(\boldsymbol{K}\big[\mathcal{I}_{il}(1)\big] + \boldsymbol{\kappa}_i\Big) - \phi\Big(\boldsymbol{K}\big[\mathcal{I}_{il}(1)\big] + \boldsymbol{\kappa}_l\Big) \\ \phi\Big(\boldsymbol{K}\big[\mathcal{I}_{il}(2)\big] + \boldsymbol{\kappa}_i\Big) - \phi\Big(\boldsymbol{K}\big[\mathcal{I}_{il}(2)\big] + \boldsymbol{\kappa}_l\Big) \\ \vdots \\ \phi\Big(\boldsymbol{K}\big[\mathcal{I}_{il}(|\mathcal{I}_{il}|)\big] + \boldsymbol{\kappa}_i\Big) - \phi\Big(\boldsymbol{K}\big[\mathcal{I}_{il}(|\mathcal{I}_{il}|)\big] + \boldsymbol{\kappa}_l\Big) \end{bmatrix}. \tag{2.42}$$

Let $\hat{\mathcal{S}}'_i\big[\mathcal{I}_{il}\big]$ ($i = 1, 2, \cdots, n$) be a $\mathbb{R}^{|\mathcal{I}_{il}|\times 1}$ vector which is defined as

$$\hat{\mathcal{S}}'_i\big[\mathcal{I}_{il}\big] := \Big[\hat{\mathcal{S}}'_i\big[\mathcal{I}_{il}(1)\big], \quad \hat{\mathcal{S}}'_i\big[\mathcal{I}_{il}(2)\big], \quad \cdots, \quad \hat{\mathcal{S}}'_i\big[\mathcal{I}_{il}(|\mathcal{I}_{il}|)\big]\Big]^T, \tag{2.43}$$

where T in the superscript denotes ordinary transpose. We can then simplify the expression of $\mathbf{D}_{il}\boldsymbol{\varphi}$

$$\mathbf{D}_{il}\boldsymbol{\varphi} = \arg\Big(\hat{\mathcal{S}}'_i\big[\mathcal{I}_{il}\big]\Big) - \arg\Big(\hat{\mathcal{S}}'_l\big[\mathcal{I}_{il}\big]\Big) + \big[\phi(\mathbf{k}_i) - \phi(\mathbf{k}_l)\big]. \tag{2.44}$$

Note $\phi(\mathbf{k}_i) - \phi(\mathbf{k}_l)$ is a constant. To count for this constant, we define an offset operator $\mathbf{D}_{il}^0 \in \mathbb{R}^{|\mathcal{I}_{il}| \times N^2}$, where we assign 1 to place that corresponds to $\mathbf{k}_i$ and -1 to $\mathbf{k}_l$. That is,

$$\mathbf{D}_{il}^0(i_1, i_2) = \begin{cases} 1, & \text{if } \boldsymbol{K}(i_2) = \boldsymbol{\kappa}_i \\ -1, & \text{if } \boldsymbol{K}(i_2) = \boldsymbol{\kappa}_l \\ 0, & \text{otherwise.} \end{cases} \tag{2.45}$$

When acting on $\boldsymbol{\varphi}$, this offset operator gives

$$\mathbf{D}_{il}^0 \boldsymbol{\varphi} = \begin{bmatrix} \phi(\mathbf{k}_i) - \phi(\mathbf{k}_l) \\ \phi(\mathbf{k}_i) - \phi(\mathbf{k}_l) \\ \vdots \\ \phi(\mathbf{k}_i) - \phi(\mathbf{k}_l) \end{bmatrix}. \tag{2.46}$$

Thus, we can use $\mathbf{D}_{il}$ and $\mathbf{D}_{il}^0$ to express the total phase difference between $\hat{\mathcal{S}}_i'$ and $\hat{\mathcal{S}}_l'$, which gives

$$(\mathbf{D}_{il} - \mathbf{D}_{il}^0)\boldsymbol{\varphi} = \arg\left(\hat{\mathcal{S}}_i'[\mathcal{I}_{il}]\right) - \arg\left(\hat{\mathcal{S}}_l'[\mathcal{I}_{il}]\right). \tag{2.47}$$

For different pairs of spectrums with overlap, we concatenate those equations, which yields

$$\mathbf{D}\boldsymbol{\varphi} = \boldsymbol{\beta}_\Delta,$$

where

$$\mathbf{D} := \begin{bmatrix} \mathbf{D}_{i_1 l_1} - \mathbf{D}_{i_1 l_1}^0 \\ \mathbf{D}_{i_2 l_2} - \mathbf{D}_{i_2 l_2}^0 \\ \vdots \\ \mathbf{D}_{i_m l_m} - \mathbf{D}_{i_m l_m}^0 \end{bmatrix} \quad \text{and} \quad \boldsymbol{\beta}_\Delta := \arg\begin{pmatrix} \hat{\mathcal{S}}_{i_1}'[\mathcal{I}_{i_1 l_1}] \\ \hat{\mathcal{S}}_{i_2}'[\mathcal{I}_{i_2 l_2}] \\ \vdots \\ \hat{\mathcal{S}}_{i_m}'[\mathcal{I}_{i_m l_m}] \end{pmatrix} - \arg\begin{pmatrix} \hat{\mathcal{S}}_{l_1}'[\mathcal{I}_{i_1 l_1}] \\ \hat{\mathcal{S}}_{l_2}'[\mathcal{I}_{i_2 l_2}] \\ \vdots \\ \hat{\mathcal{S}}_{l_m}'[\mathcal{I}_{i_m l_m}] \end{pmatrix}. \tag{2.48}$$

Instead of directly solving for the aberration term, we use Zernike polynomial to represent the aberration of the system. Thus, we introduce the Zernike operator $\mathbf{Z} \in \mathbb{R}^{N^2 \times z}$, where z is the number of Zernike coefficients to be reconstructed and each column of $\mathbf{Z}$ represents a particular Zernike mode. With Zernike decomposition, the aberration of the imaging system is given by

$$\boldsymbol{\varphi} = \mathbf{Z}\mathbf{c}, \tag{2.49}$$

where $\mathbf{c} \in \mathbb{R}^{z \times 1}$ is the corresponding Zernike coefficient. Using Zernike decomposition, we can rewrite Eq. 2.48 as

$$\boldsymbol{\beta}_\Delta = \mathbf{D}\boldsymbol{\varphi} = \mathbf{D}\mathbf{Z}\mathbf{c}. \tag{2.50}$$

We can solve the above linear equation (or the associated normal equation) to get the analytical solution of the Zernike coefficient $\mathbf{c}$. The 2D aberration $\boldsymbol{\phi}$ of the imaging system is then given by

$$\boldsymbol{\phi} = \mathrm{Flat}_{N,N}^{-1}(\boldsymbol{\varphi}) = \mathrm{Flat}_{N,N}^{-1}(\mathbf{Zc}). \tag{2.51}$$

In reality, some spatial frequencies in the spectrum can be stronger than other frequencies. It is natural to emphasize on those places, as they have a higher signal-to-noise ratio (SNR). Thus, we use a weight matrix $\mathbf{W}$ to emphasize places with high SNR. This gives

$$(\mathbf{WDZ})\mathbf{c} = \mathbf{W}\boldsymbol{\beta}_{\Delta}. \tag{2.52}$$

In our experiment, we use logarithm of the modulus of the product $\hat{S}_i(\mathbf{k})\hat{S}_l(\mathbf{k})$ as the weight matrix. Another important thing to note is that the phase difference of two spectrums might exceed 2π. Thus, we first unwrap their phase differences [21] and then solve Eq. 2.52 to extract the aberration of our imaging system.

Reconstruction using darkfield measurements

Since aberrations of our imaging system are fully determined, we can remove the contribution of the aberration term from the reconstructed spectrums. Recall that when discussing about the NA matching measurements, we have reconstructed the (modified) sampled spectrums (Eqs. 2.32 and 2.33) under NA-matching angle illumination. This means the following is known for all $|\mathbf{k}_i| = \mathrm{NA}$

$$\hat{S}_i'(\mathbf{k})e^{-j\phi(\mathbf{k}_i)} = \hat{O}(\mathbf{k})H(\mathbf{k}+\mathbf{k}_i)e^{-j\phi(\mathbf{k}_i)} = \hat{O}(\mathbf{k})\,\mathrm{Circ}_{\mathrm{NA}}(\mathbf{k}+\mathbf{k}_i)e^{j\phi(\mathbf{k}+\mathbf{k}_i)-j\phi(\mathbf{k}_i)}. \tag{2.53}$$

As the last phase factor $e^{j\phi(\mathbf{k}+\mathbf{k}_i)-j\phi(\mathbf{k}_i)}$ entirely depends on the aberration, we can now correct it and extract a clean sample spectrum, which yields

$$\hat{O}(\mathbf{k})\,\mathrm{Circ}_{\mathrm{NA}}(\mathbf{k}+\mathbf{k}_i). \tag{2.54}$$

It basically means that we recovered a piece of sample spectrum region covered by the CTF support for each measurement. Thus, these reconstructed regions can be stitched for a larger coverage in the spatial frequency domain. That is, we can gradually expand the coverage of the reconstructed $\hat{O}(\mathbf{k})$ in the spatial frequency domain. Let us define the sampled region $\mathcal{M}_m$ that denotes the spectrum covered by the first mth measurements

$$\mathcal{M}_m := \left\{\mathbf{k} \in \mathbb{R}^2 \middle| \exists i \in \{1, 2, \cdots, m\}, \mathrm{Circ}_{\mathrm{NA}}(\mathbf{k}+\mathbf{k}_i) \neq 0\right\}. \tag{2.55}$$

We also define the mask M_m which denotes the effective sampling mask for the first mth measurements

$$M_m(\mathbf{k}) := \mathbb{1}(\mathbf{k} \in \mathcal{M}_m) = \begin{cases} 1, & \text{if } \mathbf{k} \in \mathcal{M}_m, \\ 0, & \text{otherwise.} \end{cases} \tag{2.56}$$

Assume we can reconstruct the complex spectrum for every measurement (we will later show this is feasible), the reconstructed complex sample spectrum $\hat{\mathsf{R}}_m(\mathbf{k})$ using the first mth measurements can be expressed by

$$\hat{\mathsf{R}}_m(\mathbf{k}) := \hat{O}(\mathbf{k})M_m(\mathbf{k}). \tag{2.57}$$

It is worth noting that, after aberration correction, we have obtained the reconstructed spectrum, $\hat{\mathsf{R}}_{n_0}(\mathbf{k}) = \hat{O}(\mathbf{k})M_{n_0}(\mathbf{k})$, using NA-matching measurements, $i = 1, 2, \cdots, n_0$. For $i > n_0$, the corresponding illumination angle exceeds the maximal acceptance angle of the imaging system and we measure the darkfield (see Eq. 2.9 for the definition of n_0).

We note that the logarithm transformation technique used for the NA matching measurements in previous subsection fails as the darkfield may probably contain null point and the Taylor expansion might not converge. As a consequence, we need to develop another algorithm to reconstruct the sampled complex spectrum.

In this section, we show that if the sampled spectrum at ith illumination $\hat{S}'_i$ ($i > n_0$) consists of previously reconstructed (*a priori*) spectrum and other unknown part, the unknown part can be reconstructed using the known spectrum.

Let us decompose the sampled spectrum $\hat{S}'_i$ into the unknown and known part, assuming all previous measurements are reconstructed, which means $\hat{\mathsf{R}}_{i-1}(\mathbf{k}) = \hat{O}(\mathbf{k})M_{i-1}(\mathbf{k})$ is known. The known spectrum $\hat{\mathcal{P}}_i(\mathbf{k})$ at this substep is given by

$$\hat{\mathcal{P}}_i(\mathbf{k}) = \hat{\mathsf{R}}_{i-1}(\mathbf{k})H(\mathbf{k} + \mathbf{k}_i) = \hat{O}(\mathbf{k})H(\mathbf{k} + \mathbf{k}_i)M_{i-1}(\mathbf{k}). \tag{2.58}$$

Then, the unknown part $\hat{U}_i(\mathbf{k})$ is

$$\begin{aligned} \hat{U}_i(\mathbf{k}) &= \hat{S}'_i(\mathbf{k}) - \hat{\mathcal{P}}_i(\mathbf{k}) \\ &= \hat{O}(\mathbf{k})H(\mathbf{k} + \mathbf{k}_i) - \hat{O}(\mathbf{k})H(\mathbf{k} + \mathbf{k}_i)M_{i-1}(\mathbf{k}) \\ &= \hat{O}(\mathbf{k})H(\mathbf{k} + \mathbf{k}_i)\left[1 - M_{i-1}(\mathbf{k})\right]. \end{aligned} \tag{2.59}$$

Let $\mathrm{Supp}(f)$ be the support of function $f : \mathbb{R}^2 \to \mathbb{C}$, which is a set given by

$$\mathrm{Supp}(f) := \{\mathbf{x} \in \mathbb{R}^2 \big| f(\mathbf{x}) \neq 0\}. \tag{2.60}$$

By construction of $\hat{\mathcal{P}}_i(\mathbf{k})$ and $\hat{U}_i(\mathbf{k})$, it is easy to see that $\hat{\mathcal{P}}_i(\mathbf{k}) \neq 0$ suggests $\hat{U}_i(\mathbf{k}) = 0$, and vice versa. Thus, $\hat{\mathcal{P}}_i(\mathbf{k})$ and $\hat{U}_i(\mathbf{k})$ have disjoint support. Based on Eqs. 2.5 and 2.6, the measured intensity of one darkfield measurement can be expressed as

$$\begin{aligned} I_i(\mathbf{x}) &= \left|\left[\mathcal{F}^{-1}(\hat{S}_i')\right](\mathbf{x})\right|^2 = \left|\left[\mathcal{F}^{-1}(\hat{U}_i + \hat{\mathcal{P}}_i)\right](\mathbf{x})\right|^2 = \left|U_i(\mathbf{x}) + \mathcal{P}_i(\mathbf{x})\right|^2 \\ &= U_i(\mathbf{x})U_i^*(\mathbf{x}) + U_i(\mathbf{x})\mathcal{P}_i^*(\mathbf{x}) + \mathcal{P}_i(\mathbf{x})U_i^*(\mathbf{x}) + \mathcal{P}_i(\mathbf{x})\mathcal{P}_i^*(\mathbf{x}), \end{aligned} \tag{2.61}$$

where $\mathcal{P}(\mathbf{x}) := \left[\mathcal{F}^{-1}(\hat{\mathcal{P}}_i\right](\mathbf{x})$ and $U(\mathbf{x}) := \left[\mathcal{F}^{-1}(\hat{U}_i\right](\mathbf{x})$ are the known and unknown fields, respectively. Using the property of Fourier transform, we can write the Fourier transform of I_i as

$$\left[\mathcal{F}(I_i)\right](\mathbf{k}) = \left[\hat{U}_i \star \hat{U}_i\right](\mathbf{k}) + \left[\hat{U}_i \star \hat{\mathcal{P}}_i\right](\mathbf{k}) + \left[\hat{\mathcal{P}}_i \star \hat{U}_i\right](\mathbf{k}) + \left[\hat{\mathcal{P}}_i \star \hat{\mathcal{P}}_i\right](\mathbf{k}), \tag{2.62}$$

where $\star$ denotes correlation. As the known spectrum is *a priori*, we can subtract its auto-correlation from the Fourier transform of the intensity measurement. This yields

$$\left[\mathcal{F}(I_i)\right](\mathbf{k}) - \left[\hat{\mathcal{P}}_i \star \hat{\mathcal{P}}_i\right](\mathbf{k}) = \left[\hat{U}_i \star \hat{U}_i\right](\mathbf{k}) + \left[\hat{U}_i \star \hat{\mathcal{P}}_i\right](\mathbf{k}) + \left[\hat{\mathcal{P}}_i \star \hat{U}_i\right](\mathbf{k}). \tag{2.63}$$

We can see that the two cross-terms are linear with respect to $\hat{U}_i$ as correlation is a linear operator. If we consider one cross term, it naturally leads to a linear equation with respect to $\hat{U}_i$, which can be solved analytically. However, those three remaining terms cannot be easily separated due to the existence of our desired unknown part.

In the remaining part of this section, we see that the above three terms have different supports, as depicted in Fig. 2.10. Therefore, the non-overlapping part can be isolated, which allows us to construct a linear equation with respect to $\hat{U}_i$.

In experiments, the same camera is used to acquire both darkfield and NA-matching measurements. Each intensity measurement is a matrix of fixed size. Similar to the FPM, the matrix representation of the original sample's spectrum R_m is nevertheless a much larger matrix [3, 4]. The physical picture is that the tilt illumination translates the sample spectrum in the spatial frequency domain, as the initial forward model Eq. 2.1 suggests, so that different spatial frequencies of sample spectrum is moved into this smaller measurement related grid and then gets sampled. This is equivalent to cropping out a specific region of the sample spectrum and moving the cropped part to the center when applying the pupil translation model in Eq. 2.6. In our following discussion, we assume the matrix representations of both the unknown and the known spectrum are centered, thus having the same dimension as the measurement.

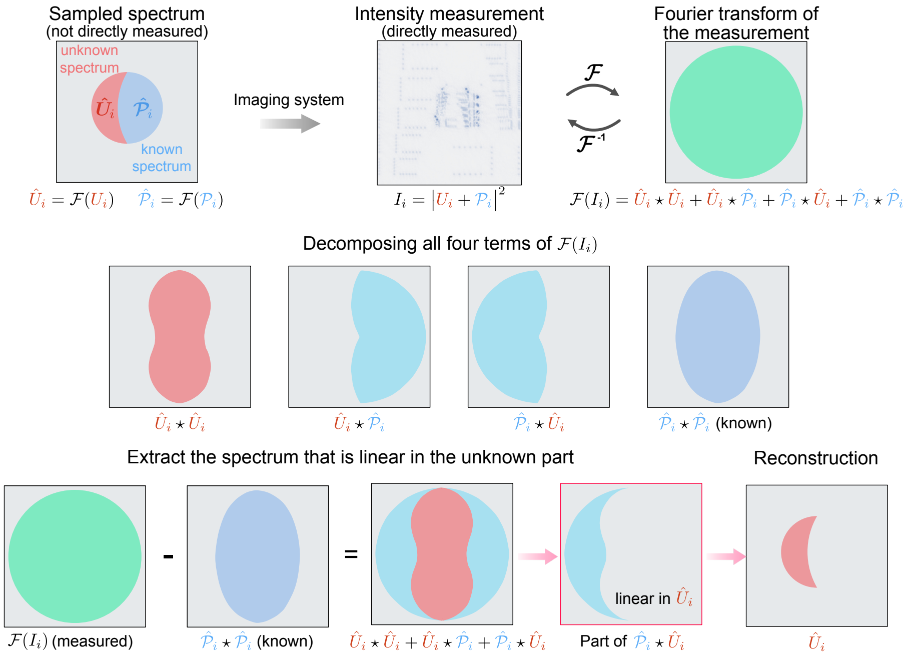

Figure 2.10: Complex field reconstruction using darkfield measurements. In this visualization, the sampled spectrum is centered, which is basically a cropped version of the spectrum translation model (Eq. 2.6). The sampled spectrum is decomposed into two disjoint region, namely the unknown and known spectrum. As the spectrum of the known part is given, the intensity of the corresponding field can be calculated. When subtracting the Fourier transform of the measured intensity and the Fourier transform of this calculated intensity, three term are left: the auto-correlation of the unknown part, and two cross-correlations of the known and unknown part. For one of the cross-correlation terms that does not overlap with the other two, we obtain a linear equation with respect to the unknown part. Thus, we can solve this linear equation for the closed-form solution of the unknown spectrum. That is, the complex spectrum can be reconstructed by incorporating the known information.

We focus on $\left[\hat{\mathcal{P}}_i \star \hat{U}_i\right](\mathbf{k})$ and let Q_i be the non-intersecting set, which is defined as

$$Q_i := \left\{\mathbf{k} \in \mathbb{R}^2 \middle| \mathbf{k} \in \mathrm{Supp}(\hat{\mathcal{P}}_i \star \hat{U}_i) \setminus \left[\mathrm{Supp}(\hat{U}_i \star \hat{\mathcal{P}}_i) \cup \mathrm{Supp}(\hat{U}_i \star \hat{U}_i)\right]\right\}. \tag{2.64}$$

By construction, we have

$$\left[\mathcal{F}(I_i)\right](\mathbf{k}) - \left[\hat{\mathcal{P}}_i \star \hat{\mathcal{P}}_i\right](\mathbf{k}) = \left[\hat{\mathcal{P}}_i \star \hat{U}_i\right](\mathbf{k}), \quad \forall \mathbf{k} \in Q_i. \tag{2.65}$$

We define this masked subtraction as L_i, that is

$$\begin{aligned} L_i(\mathbf{k}) &:= \Big(\left[\mathcal{F}(I_i)\right](\mathbf{k}) - \left[\hat{\mathcal{P}}_i \star \hat{\mathcal{P}}_i\right](\mathbf{k})\Big)\mathbb{1}(\mathbf{k} \in Q_i) \\ &= \begin{cases} \left[\hat{\mathcal{P}}_i \star \hat{U}_i\right](\mathbf{k}), & \text{if } \mathbf{k} \in Q_i, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \tag{2.66}$$

Let $\hat{\mathbf{U}}_i \in \mathbb{C}^{N\times N}$ be the matrix version of the (centered) unknown part and $\hat{\boldsymbol{\mathcal{U}}}_i \in \mathbb{C}^{N^2\times 1}$ be its flattened vector $\hat{\boldsymbol{\mathcal{U}}}_i = \mathrm{Flat}_{N,N}(\hat{\mathbf{U}}_i)$, $\mathbf{L}_i \in \mathbb{C}^{N\times N}$ be the (centered) matrix version of L_i and $\boldsymbol{\mathcal{L}}_i \in \mathbb{C}^{N^2\times 1}$ be its flattened vector $\boldsymbol{\mathcal{L}}_i = \mathrm{Flat}_{N,N}(\mathbf{L}_i)$, and $\hat{\boldsymbol{\mathcal{P}}}_i$ be the (centered) matrix version of the known part. In general, the correlation of two matrices of size $N \times N$ would be a matrix of size $(2N-1)\times(2N-1)$. Based on the Nyquist theorem, the nonzero part of $\hat{\mathbf{U}}_i + \hat{\boldsymbol{\mathcal{P}}}_i$ is contained in a $\frac{N}{2}\times\frac{N}{2}$ box. As a consequence, we can calculate the correlation of $\hat{\mathbf{U}}_i$ and $\hat{\boldsymbol{\mathcal{P}}}_i$ using a $N\times N$ grid.

Let us construct a (sparse) correlation operator $\mathbf{C}_i$ that takes all nonzero elements in the unknown spectrum $\hat{\boldsymbol{\mathcal{U}}}_i$ and gives $\boldsymbol{\mathcal{L}}_i$. We first focus on the t_1th row and t_2th column of $\mathbf{L}_i$, which corresponds to mth element of vector $\boldsymbol{\mathcal{L}}_i$, where $m = t_1 + N(t_2-1)$ and $t_1, t_2 \in \{1, 2, \cdots, N\}$. Let $t_1' = t_1 - c_0$ and $t_2' = t_2 - c_0$. For simplicity, we also write $\upsilon_\zeta^{\mathrm{lower}} = \max\{1, 1+t_\zeta'\}$ and $\upsilon_\zeta^{\mathrm{upper}} = \min\{N, N+t_\zeta'\}$ for $\zeta \in \{1,2\}$. Then, we have (recall that in Eq. 2.39 we define $\boldsymbol{K}$ as the grid version of the k vector and also $\mathbf{K}_u(c_0,c_0) = \mathbf{K}_v(c_0,c_0) = 0$)

$$
\begin{aligned}
\boldsymbol{\mathcal{L}}_i(m) &= \begin{cases} \left[\hat{\boldsymbol{\mathcal{P}}}_i \star \hat{\mathbf{U}}_i\right]\left[\boldsymbol{K}(m)\right] & \text{if } \boldsymbol{K}(m) \in \boldsymbol{Q}_i \\ 0, & \text{otherwise} \end{cases} \\
&= \begin{cases} \displaystyle\sum_{i_1=\upsilon_1^{\mathrm{lower}}}^{\upsilon_1^{\mathrm{upper}}} \sum_{i_2=\upsilon_2^{\mathrm{lower}}}^{\upsilon_2^{\mathrm{upper}}} \hat{\boldsymbol{\mathcal{P}}}_i^*(i_1 - t_1', i_2 - t_2')\hat{\mathbf{U}}_i(i_1, i_2), & \text{if } \boldsymbol{K}(m) \in \boldsymbol{Q}_i, \\ 0, & \text{otherwise.} \end{cases}
\end{aligned} \tag{2.67}
$$

For $\boldsymbol{K}(m) \in \boldsymbol{Q}_i$, we can construct a matrix $\mathbf{G}_i^m \in \mathbb{C}^{N\times N}$ that is defined as

$$
\mathbf{G}_i^m(i_1, i_2) := \begin{cases} \hat{\boldsymbol{\mathcal{P}}}_i^*(i_1 - t_1', i_2 - t_2'), & \text{if } \upsilon_\zeta^{\mathrm{lower}} \le i_\zeta \le \upsilon_\zeta^{\mathrm{upper}}, \zeta \in \{1,2\} \\ 0, & \text{otherwise.} \end{cases} \tag{2.68}
$$

With this definition, we have

$$
\boldsymbol{\mathcal{L}}_i(m) = \begin{cases} \left[\,\mathrm{Flat}_{N,N}(\mathbf{G}_i^m)\right]^T \hat{\boldsymbol{\mathcal{U}}}_i, & \text{if } \boldsymbol{K}(m) \in \boldsymbol{Q}_i \\ 0, & \text{otherwise.} \end{cases} \tag{2.69}
$$

We define an index set $\mathfrak{L}_i$ that denotes this special region that is linear in $\hat{\boldsymbol{\mathcal{U}}}_i$

$$
\mathfrak{L}_i := \{m = 1, 2, \cdots, N^2 \big| \boldsymbol{K}(m) \in \boldsymbol{Q}_i\}. \tag{2.70}
$$

As we have done before, we abuse the notation so that $\mathfrak{L}_i(m)$ indicates the mth smallest element in $\mathfrak{L}_i$ and define $\boldsymbol{\mathcal{L}}_i(\mathfrak{L}_i)$ as

$$
\boldsymbol{\mathcal{L}}_i(\mathfrak{L}_i) := \Big[\boldsymbol{\mathcal{L}}_i\big[\mathfrak{L}_i(1)\big], \quad \boldsymbol{\mathcal{L}}_i\big[\mathfrak{L}_i(2)\big], \quad \cdots, \quad \boldsymbol{\mathcal{L}}_i\big[\mathfrak{L}_i(|\mathfrak{L}_i|)\big]\Big]^T, \tag{2.71}
$$

where $|\mathfrak{L}_i|$ is the cardinal of $\mathfrak{L}_i$. We can then construct a correlation operator. Let $\mathbf{C}_i^F \in \mathbb{C}^{|\mathfrak{L}_i| \times N^2}$ be a $|\mathfrak{L}_i| \times N^2$ matrix and let its mth row $\mathbf{C}_i^F(m, \cdot)$ be

$$\mathbf{C}_i^F(m, \cdot) = \left[\mathrm{Flat}_{N,N}(\mathbf{G}_i^{\mathfrak{L}_i(m)})\right]^T, \tag{2.72}$$

we then have

$$\mathbf{C}_i^F \hat{\boldsymbol{\mathcal{U}}}_i = \boldsymbol{\mathcal{L}}_i(\mathfrak{L}_i). \tag{2.73}$$

Graphically, the construction of the correlation operator $\mathbf{C}_i^F$ is shown in Fig. 2.11.

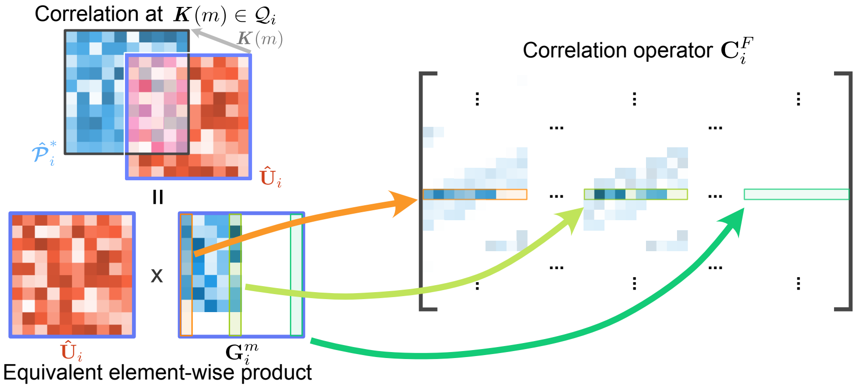

Figure 2.11: Construction of the correlation operator $\mathbf{C}_i^F$. For vector $\boldsymbol{K}(m)$ that admits a measurement that is linear in the unknown spectrum, we construct an associated matrix $\mathbf{G}_i^m$. This matrix is constructed such that the correlation of the unknown and known part at $\boldsymbol{K}(m)$ equals $\sum_{i_1,i_2=1}^{N} \hat{\mathrm{U}}_i(i_1, i_2)\mathbf{G}_i^m(i_1, i_2)$, the summation of the element-wise multiplication of this matrix and the unknown spectrum. The flattened vector of $\mathbf{G}_i^m$ serves as one row of the correlation matrix $\mathbf{C}_i^F$, as illustrated in the figure.

By construction of the unknown spectrum $\hat{U}_i$, we know that the only locations that it can be nonzero are where the following holds

$$H(\mathbf{k} + \mathbf{k}_i)\left[1 - M_{i-1}(\mathbf{k})\right] \neq 0. \tag{2.74}$$

Using the above equation, we can easily find the corresponding nonzero elements in the flattened vector $\hat{\boldsymbol{\mathcal{U}}}_i$. We define an index set $\mathbf{N}_i$ consists of the indices of these nonzero elements

$$\mathbf{N}_i := \left\{ m = 1, 2, \cdots, N^2 \middle| H\left[\boldsymbol{K}(m) + \boldsymbol{\kappa}_i\right]\left(1 - M_{i-1}\left[\boldsymbol{K}(m)\right]\right) \neq 0 \right\}, \tag{2.75}$$

and let $\mathbf{N}_i(m)$ be the mth smallest element in $\mathbf{N}_i$, and $|\mathbf{N}_i|$ be the cardinal of $\mathbf{N}_i$. To construct the correlation operator $\mathbf{C}_i$ that encodes the sparsity of the unknown spectrum, we simply keep all columns of $\mathbf{C}_i^F$ whose indices belong to set $\mathbf{N}_i$ and throw away all other columns. This gives the definition of $\mathbf{C}_i \in \mathbb{C}^{|\mathfrak{L}_i| \times |\mathbf{N}_i|}$

$$\mathbf{C}_i(i_1, m) := \mathbf{C}_i^F\left[i_1, \mathbf{N}_i(m)\right], \quad i_1 = 1, 2, \cdots, |\mathfrak{L}_i| \text{ and } m = 1, 2, \cdots, |\mathbf{N}_i|. \tag{2.76}$$

Then, we have the following linear equation with respect to the nonzero elements of the unknown spectrum

$$\mathbf{C}_i \hat{\mathcal{U}}_i(\mathbf{N}_i) = \mathcal{L}_i(\mathfrak{L}_i), \tag{2.77}$$

where $\hat{\mathcal{U}}_i(\mathbf{N}_i) := \left[\hat{\mathcal{U}}_i\left[\mathbf{N}_i(1)\right], \quad \hat{\mathcal{U}}_i\left[\mathbf{N}_i(2)\right], \quad \cdots, \quad \hat{\mathcal{U}}_i\left[\mathbf{N}_i(|\mathbf{N}_i|)\right]\right]^T$.

To solve this equation, we require the rank of matrix $\mathbf{C}_i$ to be at least $|\mathbf{N}_i|$. We can show this can be satisfied if the known spectrum covers the semicircle of the circular CTF. As shown in Fig. 2.12, the autocorrelation of a semicircle is around 4 times larger than itself. If we assume the CTF is of radius r_0, and the known specturm is a semicircle, the area of the unknown spectrum is then $\frac{1}{2}\pi r_0^2$. For a circle (area is πr_0^2), its autocorrelation is strictly 4 times larger in size (area is $4\pi r_0^2$). Thus, the linear region has an area of

$$\text{Area}(\mathcal{Q}_i) = \frac{1}{2}\left[4\pi r_0^2 - (4r_0^2 + \pi r_0^2)\right] = \frac{3\pi r_0^2 - 4r_0^2}{2}, \tag{2.78}$$

which is approximately 1.7 times larger than the area of the unknown part. That is, if the known spectrum occupies 50% of the spectrum, the rank of matrix $\mathbf{C}_i$ can be well above $|\mathbf{N}_i|$. Numerically, we find a safe choice is to let the unknown spectrum occupy over 42% of the measured spectrum, assume the CTF is circular.

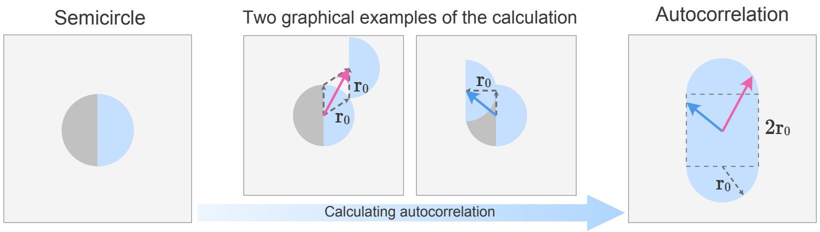

Figure 2.12: Autocorrelation of a semicircle with radius r_0. The shape of the autocorreltion result has an area of $4r_0^2 + \pi r_0^2$ which is around 4.5 times larger than the area enclosed by the semicircle.

By solving Eq. 2.77, we obtain the closed-form solution of the unknown spectrum. That is, we reconstruct the following spectrum

$$\hat{U}_i(\mathbf{k}) = \hat{O}(\mathbf{k}) H(\mathbf{k} + \mathbf{k}_i)\left[1 - M_{i-1}(\mathbf{k})\right]. \tag{2.79}$$

As the aberration of the system is determined, we can correct for the aberration, which gives the aberration corrected spectrum $\hat{O}(\mathbf{k})\,\mathrm{Circ}_{\mathrm{NA}}(\mathbf{k}+\mathbf{k}_i)\left[1-M_{i-1}(\mathbf{k})\right]$. Because the intensity of the field is directly measured in the darkfield measurement and the phase is unknown, we use the square root of this measured intensity as the modulus of the complex field for maintaining the (point-wise) energy and robustness.

Until now, we show that the complex spectrum sampled with ith illumination can be reconstructed using *a priori* knowledge of the known spectrum. We can expand the entire reconstructed spectrum by integrating this newly reconstructed spectrum. The overall reconstructed spectrum using ith measurement is given by $\hat{O}(\mathbf{k})M_{i-1}(\mathbf{k})+\hat{O}(\mathbf{k})\,\mathrm{Circ}_{\mathrm{NA}}(\mathbf{k}+\mathbf{k}_i)\left[1-M_{i-1}(\mathbf{k})\right]$, which is exactly $\hat{O}(\mathbf{k})M_i(\mathbf{k})$ by definition of $M_i(\mathbf{k})$. That is, we obtain an extended complex spectrum reconstution $\hat{\mathsf{R}}_i$ after darkfield reconstruction at the ith sub-step. This reconstructed field serves as our *a priori* knowledge in the reconstruction of the $(i+1)$th sub-step.

We note that for a dense LED array, the area of the unknown spectrum can be quite small because their effective CTFs cover similar area for two closely spaced LEDs (Eq. 2.6). In such case, we can fill in the new spectrum when it is necessary to solve the linear equation formed in Eq. 2.77 (e.g., when the system becomes undetermined if we do not fill in the new spectrum). By doing this, the spectrum reconstructions prior to the stitching process are independent and they have overlaps in the spatial frequency domain. Therefore, we can average over the overlap for improved robustness of the reconstruction algorithm.

Once the darkfield reconstruction is done for all measurements, we have reconstructed all sampled spectrums and the extended complex spectrum $\hat{\mathsf{R}}_n$, which is our high-resolution, aberration-free complex field reconstruction.

2.6 Supporting materials for APIC

In the following section, we will show some supporting materials based on both simulation and experiment of APIC.

We demonstrate an system calibration technique that can be used to optimize the illumination angles in experiments. Resolution of APIC and the full result using the red, green and blue channels are also presented in this section. To investigate the performance of APIC under different cases, we conducted multiple simulations and present those results.

System calibration

To do reconstruction using APIC, we need to know the transverse illumination vector $\mathbf{k}_i$ for each measurement, as it tells us which area of the whole sample's spectrum is measured (Eq. 2.6). This indicates that we need to determine the angle of each tilted illumination. To do that, we used the previously developed circle-finding algorithm to find the exact illumination angle for the NA-matching measurements [22]. The brightfield measurements whose illumination angles were below the acceptance angle of our imaging system were collected as well. These brightfield measurements were used for geometrically calibrating the angles associated with our darkfield measurements. We note that these brightfield measurements are only for calibration purpose and not for reconstruction in APIC.

The illumination unit consisted of a LED ring and a LED array. The LED ring was attached on top of the LED array and was used for the NA-matching measurement. This unit was mounted on a motorized transnational stage for height adjustment. We adjusted its tilt and height to exactly match the illumination angle of the ring LED and the acceptance angle of our imaging system. To find the exact height, we first moved the LED unit close to our sample such that the ring LED produced the darkfield measurement. Then, we gradually increased the separation between the LED and the sample until we saw the image under the ring LED illumination transited from darkfield to brightfield. The transition point is our desired height. Once the height and tilt of the system were fixed, we acquired all calibration data and calibrated the illumination angles for all LEDs. We also used a high NA objective to measure the relative intensity of each LED with a blank slide. The high NA objective was selected such that the incident light from any LED can directly enter the system. It needs to be emphasized that the intensity calibration is done only once with a high NA, small field-of-view objective, we acquired all our actual experiment data with a low NA, large field-of-view objective. In our experiment, we normalized the measurements using the measured relative intensities and then conducted the reconstruction in APIC.

We note that we can reconstruct the calibration data with the geometrically calculated darkfield LED illumination angle and then use the reconstructed complex field to further optimize the illumination angle by searching over a pre-defined finer grid. Once this is done, we fix the calibrated angles and use them for all other measurements. The entire process is illustrated in Fig. 2.13.

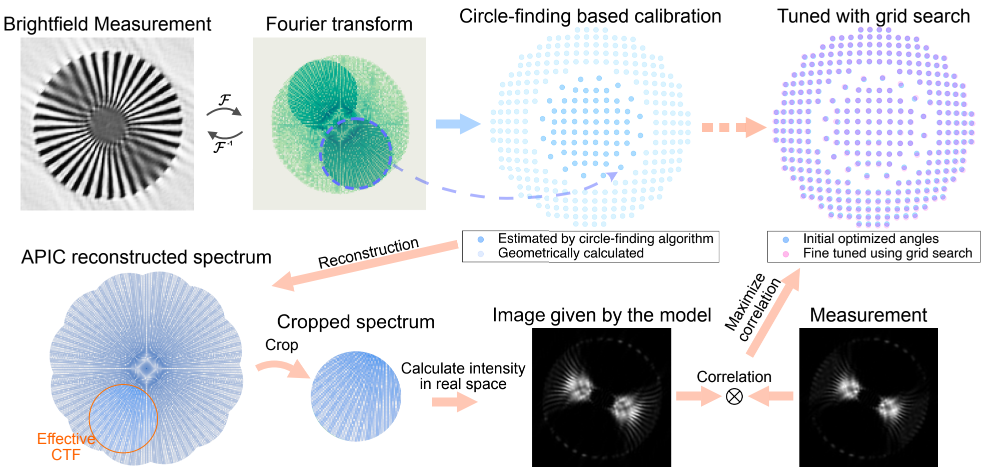

Figure 2.13: Calibrating the illumination k-vector. By locating the center of the circle in the Fourier transform of the measurement, we extract the corresponding $\mathbf{k}_i$ in the spatial frequency domain. Using the separation of the LEDs on the array and the estimated brightfield LED illumination angles, we can calculate the darkfield LED illumination angles using such geometry. With those, APIC could reconstruct sample's complex field, which can be used to further optimize the illumination angles by maximizing the correlation between the real measurement and the image obtained with the forward model. CTF: coherent transfer function.

Result of FPM and APIC using reduced dataset

In previous sections, we acquired 316 images for one sample. As we explained in previous sections, this large redundancy was chosen to show the best performance of Fourier Ptychographic Microscopy (FPM). For APIC, it is not necessary to have such a large redundancy for an accurate reconstruction. Here, we reduced the dataset so that there are 9 bright field measurements, 8 NA-matching measurements and 28 darkfield measurements in this reduced dataset. In our reconstruction, FPM used all these 45 images while APIC used 36 images (8 NA-matching measurements and 28 darkfield measurements). The reconstruction results using this reduced dataset are shown in Fig. 2.14. For comparison, we also included the reconstruction results when feeding in the entire dataset.

When FPM is not given the privilege of having a highly redundant dataset, its reconstruction result can be severely disturbed by the aberration of an imaging system whose phase variation exceeds $\frac{2\pi}{5}$ (NA of the objective: 0.25). We see that although FPM partially reconstructed the high spatial frequency information of the Siemens star target using the full dataset, it failed to maintain even the low spatial frequency information when a dataset with approximately 7 times fewer

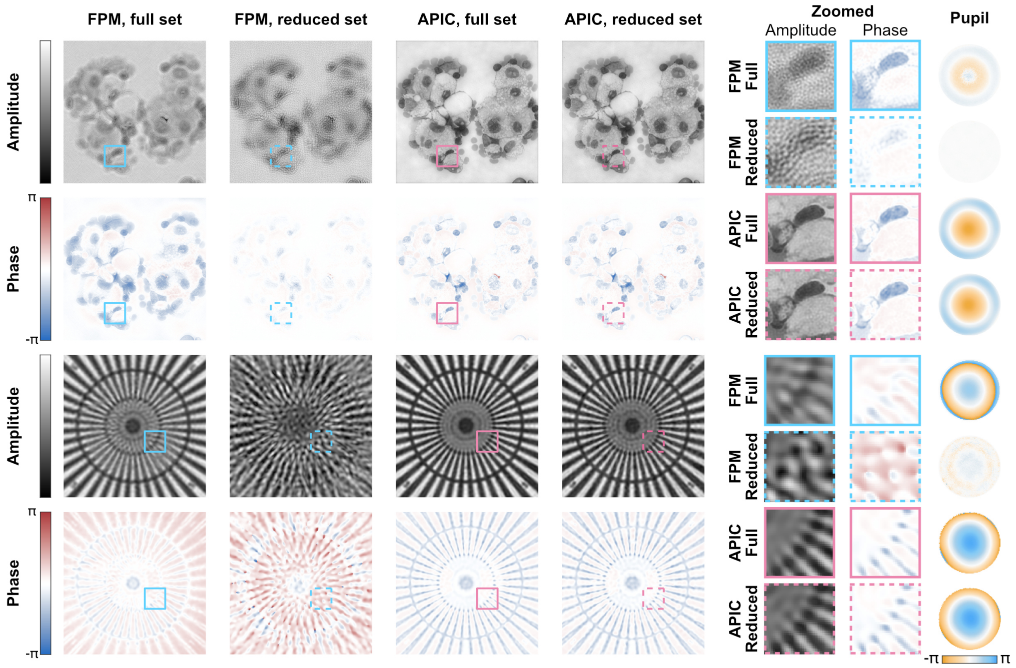

Figure 2.14: APIC and FPM using the full dataset and a reduced dataset. The full dataset is largely redundant as the overlap of two spectrum is around 86%. The nominal overlap ratio reduces to 65% for the reduced dataset. Siemens star target was defocused by 16 µm in this experiment. The zoomed images of the amplitude and phase reconstructions of APIC and FPM are shown on the right, respectively.

measurements was provided. In contrast, APIC, retrieved both the high and low spatial frequency information in either case. As we can see from Fig. 2.14, APIC generated a reliable reconstructions using both the entire dataset and the reduced dataset.

Resolution quantification

To quantify the resolution achieved in APIC, we imaged a Siemens star target. In addition, we utilized LEDs whose illumination angles were smaller than the acceptance angle of our used objective for FPM. Those are the normal brightfield measurements required for FPM. We compared the resolution achieved by APIC and FPM. Here, we performed reconstruction on the in-focus Siemens star target shown in Fig. 2.7. To compare their resolution, we first calculated the smallest radius r where at least 10% contrast was preserved for any of the two reconstructed amplitude for the Siemens star target. We then plotted the radial profiles of APIC and FPM using this radius, respectively.

From Fig. 2.15, we see that APIC and FPM have similar radial profiles. The

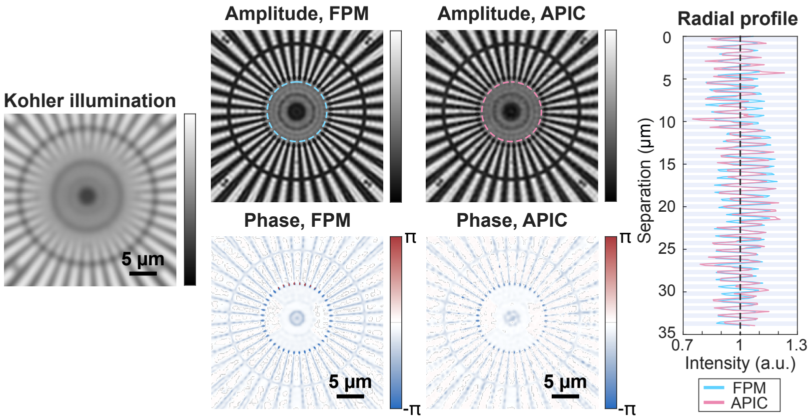

Figure 2.15: The resolution of FPM and APIC. We conducted the reconstruction using the in focus data and the reconstructed amplitude and phase of FPM and APIC are shown in the middle. Their radial profiles along the circles highlighted in the reconstructed amplitude images are shown on the right.

maximal illumination NA in this experiment is approximately 0.5 and the objective NA equals 0.25. We used red LEDs in the experiment and its central wavelength is 632 nm. Thus, the theoretical resolution of our system is 842 nm. Using the radial profiles in Fig. 2.15, the resolution we experimentally achieved for both FPM and APIC are approximately 867 nm, which is close to the theoretical resolution limit.

Reconstruction of a hematoxylin and eosin stained sample

We have included the result of the blue channel in Fig. 2.8. To see the reconstructions under different LED illumination colors, we showed the full comparison of reconstruction results of all three channels using FPM and APIC.

As we can see from Fig. 2.16, the reconstructions of FPM appeared grainy and the high spatial frequency information were only partially recovered. This indicates that FPM did not work well with this weakly absorptive sample. We also found that there were places where the sign of quantitative phase reconstructions of FPM for the green channel appeared to be flipped. In phase reconstruction shown in Fig. 2.16, some of the negative phases in the red and blue channel (phases shown in blue) become positive in the green channel (phases shown in red). The quantitative phases of APIC were much more consistent among all three channels. This indicates the aberration and complex field reconstruction of APIC is considerably more accurate compared with FPM.

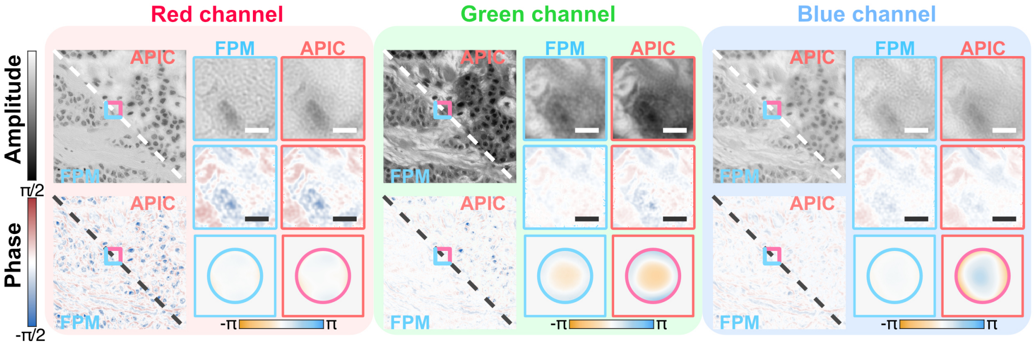

Figure 2.16: Full reconstruction, H&E stained breast cancer cells. The complex field reconstructions and retrieved aberrations of each channel are shown in the red, green and blue highlighted boxes, respectively. For each group, FPM reconstructions are shown in the lower left and APIC reconstructions are shown in the upper right. The zoomed images of the amplitude and phase reconstructions are shown on the right of each boxes. Scale bar for the zoomed images: 5 µm.

Reconstruction time

We investigated the reconstruction time taken by FPM and APIC under different patch sizes. In this simulation, the overlap ratio was set to approximately 60%. The dataset consisted of 76 darkfield measurements and 16 NA-matching measurements. We used the same convergence criterion for first and second order FPM reconstruction methods as we did in experiments and simulations. We added moderate aberration in our simulation (the phase standard deviation of the simulated aberration is approximately 0.6 radian) so that FPM could give a reasonable reconstructions. We note that FPM might run much longer when dealing with more complicated cases, as the convergence criteria become more difficult to reach in the reconstruction of FPM. The reconstructions were conducted using a personal computer with 16 GB RAM (CPU: Intel Core i5-8259U).

Figure 2.17 shows the runtime comparison of FPM and AMIM as a function of the image patch size. When the side length of the patch equals 256, these three methods needs approximately the same amount of time to complete. We note that the reconstruction time of APIC also depends on the overlap ratio of the sampled spectrums. When using the highly aberrated images of the thyroid sample (Fig. 2.7), the reconstruction time of the second order FPM and APIC was comparable when reconstructing a square patch with side length of 512 pixels. Nonetheless, we see that APIC can be much faster when a small patch size is chosen. The computation efficiency advantage of APIC for small patch size is in line with another important computational consideration, namely parallel processing. By splitting the

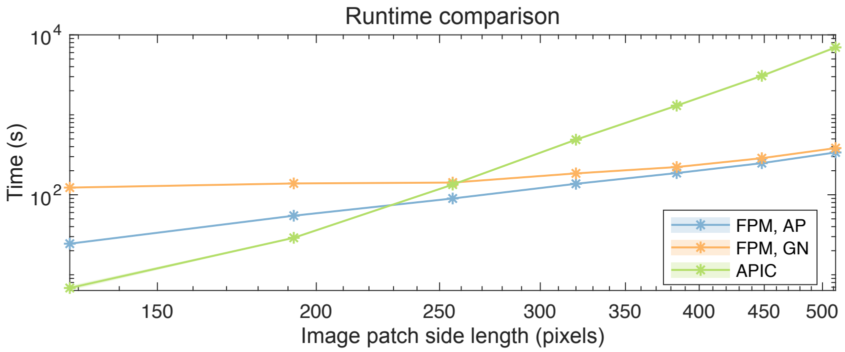

Figure 2.17: Runtime comparison of different algorithms. The patch sizes used in this simulation formed an arithmetic sequence ranging from 128 to 512 pixels, with a common difference of 64 pixels. The reconstruction of each method was repeated 5 times for each patch size, and the mean and standard deviation were calculated. The standard deviation is illustrated by the colored area behind the solid lines. In log-scale, the standard deviation is small compared with the mean value. Thus, we only see the solid lines showing up in the figure. AP denotes the original FPM reconstruction algorithm with EPRY and GN denotes the second order Gauss-Newton reconstruction algorithm.

whole image into smaller patches, more processors can be simultaneously used for reconstruction.

Result of spatial-domain Kramers-Kronig method and APIC

Here, we reconstructed the complex field using the recently proposed sptial Kramers-Kronig algorithm, which does not take complex aberration into consideration [17–19].

We used a USAF target for simulation. The LED illumination angles were assumed to be known and an aberration-free imaging system was used in our simulation. We simulated both the darkfield measurements and the NA matching measurements and fed those measurements into the spatial Kramers-Kronig method and APIC. The reconstructions of the two different methods are shown in Fig. 2.18.

As the spatial Kramers-Kronig method cannot perform reconstruction with darkfield measurements, its resolution is limited. APIC is the first of its kind that can do analytical complex field reconstruction using darkfield measurements. With the more general phase retrieval framework established in APIC, we can see that the resolution of APIC is much enhanced compared with the spatial Kramers-Kronig method.

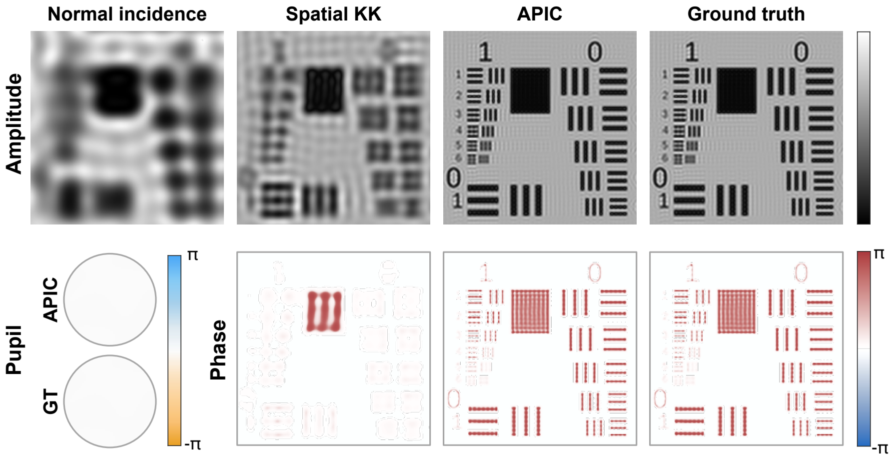

Figure 2.18: Comparison of the spatial Kramers-Kronig (spatial KK) method and APIC. The image under normal incidence is shown on the left and the ground truth of the USAF target we used in the simulation are shown on the right. The ground truth of the aberration and APIC recovered aberration are shown on the lower left of this image. In this simulation, no aberration was introduced. GT: ground truth.

Robustness of APIC with respect to aberration

We conducted simulations under different aberration levels and used APIC to reconstruct the complex field. We simulated a complex USAF target as our object and added different aberrations in our measurements. We assumed the illumination angles were known and no noise was simulated. For the NA-matching angle illumination, 16 evenly spaced LEDs were simulated. The overlap ratio of two adjacent sampled spectrums is around 60%. The ground truth was generated by cropping the complex object spectrum using the sampled region $M_n(\mathbf{k})$. In other words, the ground truth is the inverse Fourier transform of the spectrum $\hat{O}(\mathbf{k})M_n(\mathbf{k})$. To visually perceive different aberration levels, we also simulated the image that would be obtained by the aberrated imaging system under the normal incidence illumination. This normal incidence measurement was just for visualization and was not used in any APIC's reconstruction process.

The reconstruction results are shown in Fig. 2.19. From this simulation, we can see APIC is exceptionally tolerant against aberration and can accurately extract the aberration of the imaging system under extremely high aberration level (the largest phase standard deviation exceeds 1.5π). The reconstructed phase and amplitude for all cases matched up with the ground truth.

For comparison, we also used FPM to do reconstruction using two less aberrated cases above. We note that FPM needs a different set of brightfield measurements

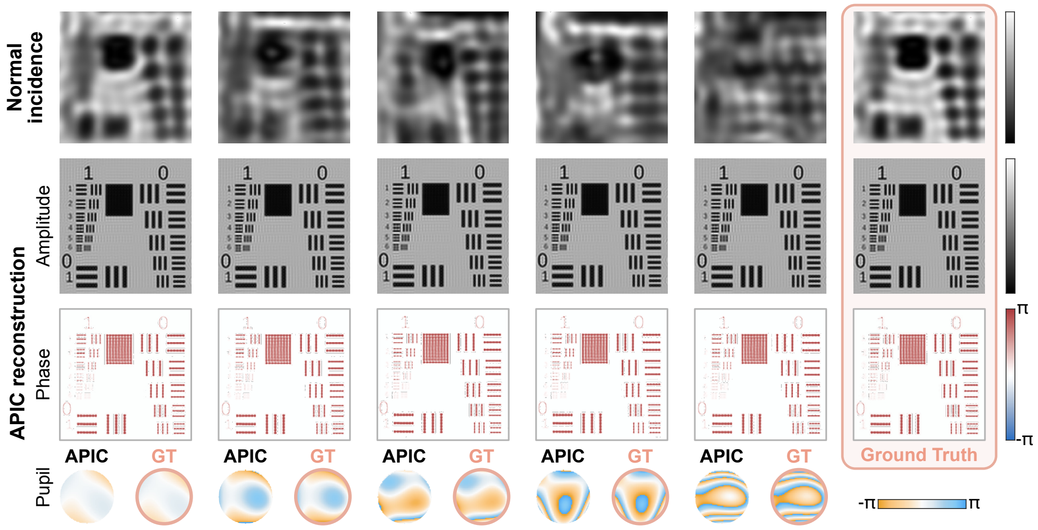

Figure 2.19: APIC under different levels of simulated aberrations. The first row shows the corresponding image under normal incidence and is aberrated. The second and third rows are the reconstructed amplitude and phase, respectively. The last row is the reconstructed aberration and the actual simulated aberration. The last column shows the ground truth of object's amplitude and phase, as well as the image under normal incidence when captured by an aberration-free imaging system.

to complete the reconstruction. As such, we decided to match the number of brightfield measurements in FPM and APIC. If there were in total n_0 NA-matching measurements in APIC, we uniformly sampled n_0 brightfield illumination angles in our simulation. The darkfield measurements were shared by FPM and APIC. We adopted two different reconstruction algorithms for FPM, namely the original alternating projection algorithm (the original Gerchberg–Saxton algorithm combined with EPRY for aberration correction) [3–5] and the second order Gauss-Newton method [7, 8]. For FPM, we conducted the reconstruction with 6 different sets of parameters. As we knew the ground truth in our simulation, we manually chose one of these parameter sets so that its corresponding result is closest to the ground truth.

From the simulation results shown in Fig. 2.20, FPM worked well with mild aberrations. When the imaging system had a relatively small aberration (its phase standard deviation is approximately 0.15π), both the original alternating projection method (implemented with EPRY for aberration correction) and the second order Gauss-Newton method successfully reconstructed the aberrations. Their amplitude reconstructions were also closely matched with the ground truth. However, we can see that the phase reconstruction of the second order method shows better correspondence to the ground truth. When the aberration in our simulation became

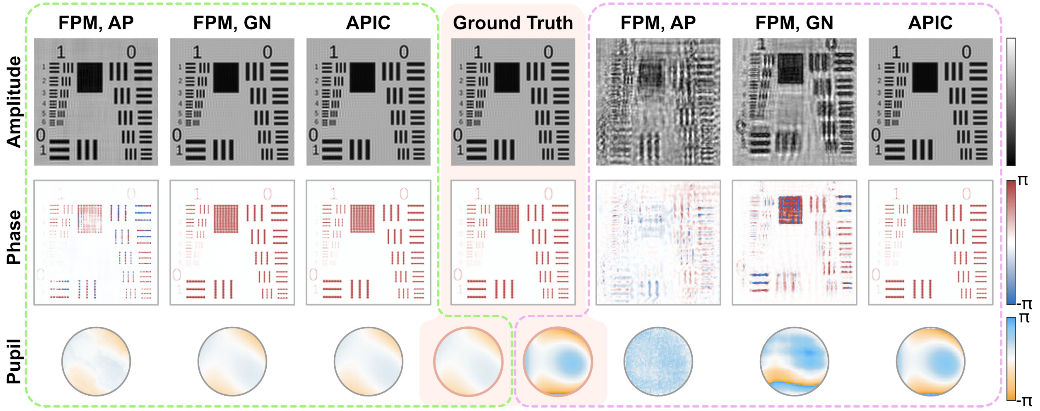

Figure 2.20: Comparison of FPM and APIC under moderate aberrations. AP denotes the alternating projection method (combined with EPRY) and GN denotes Gauss-Newton method. The ground truth of system's aberrations and the complex object are shown in the middle. In this simulation, two levels of moderate aberrations were applied. The results on the left correspond to the case where the gentler one among the two simulated aberration was used. The results on the right correspond to the more severe case.

a bit more severe (standard deviation reaches 0.4π), both FPM methods failed drastically. The second order FPM method works slightly better than the first order algorithm as it partially reconstructed the high-frequency information. However, those reconstructed images are severally distorted. APIC, on the contrary, works well with these different levels of aberrations.

Comparison under different signal-to-noise ratios

Here, we run APIC with different signal-to-noise ratios (SNRs) to see its performance under different scenarios. In this simulation, we assumed the illumination angles were known and an ideal aberration-free imaging system was used. Poisson noise was added in each measurement. The average number of photons for NA-matching measurements is shown on the top of Fig. 2.21. All other simulation parameters were the same.

If no noise was added to the measurement, we can see that APIC produced result that matched up with the ground truth. When the SNR was low, APIC became more noisy and exhibited degraded resolution. Nonetheless, it preserved the high frequency details that were not captured in the normal incidence measurement.

Then we conducted another two simulations to compare APIC with FPM under different SNRs. As before, we replaced the NA-matching measurements with n_0

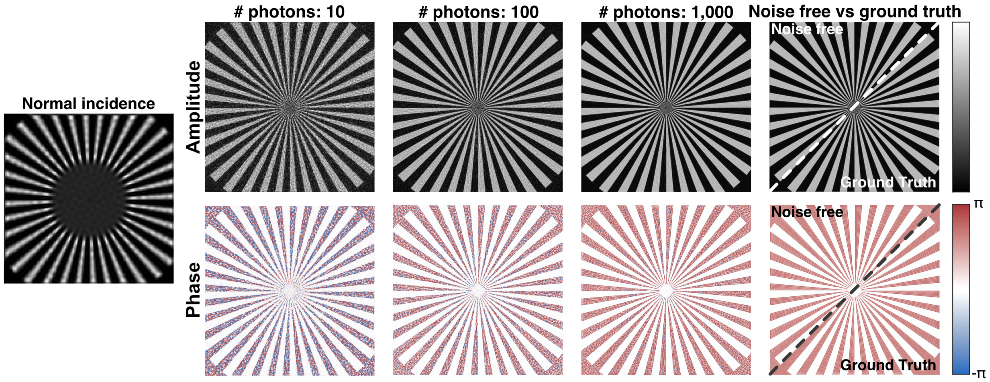

Figure 2.21: APIC under different SNRs. The measurement under normal incidence is shown on the far left, with no noise added. This measurement is for illustration purpose and is not used in APIC's reconstruction. Poisson noise was added to each measurement in this simulation. The numbers shown on the top are the average number of photons under NA-matching angle illumination. When the average photons is 10 for the NA-matching measurements, the average number of photons for all darkfield measurements is around 0.11.

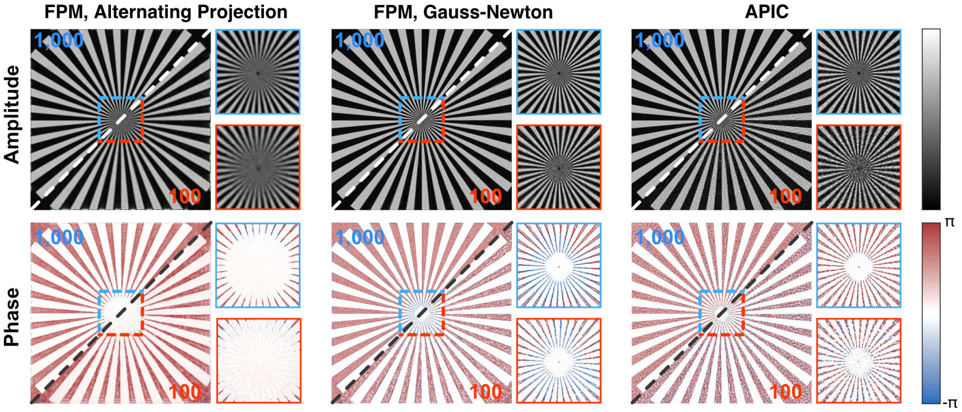

Figure 2.22: Comparison of FPM and APIC under different noise levels. For each group, the upper left panel shows the reconstruction result when the average number of photons of NA-matching measurements is 1,000, and the lower right panel shows the result when the number is 100. The zoomed version of the result within the box is shown on the far right of each group. For FPM, we performed the reconstruction using 6 different parameter sets and demonstrated the best reconstruction result we obtained.

brightfield measurements to construct the dataset for FPM. All darkfield measurements were shared by APIC and FPM. For FPM, we chose 6 different sets of parameters and selected the best results in our simulation.

In the first simulation, we simulated a complex Siemens star target and added different levels of noise to the simulated dataset. We can see from Fig. 2.22 that the second order FPM algorithm performs well under low SNR (when the average number of photons equals 100 for the NA-matching measurements). This iterative algorithm trades the reconstruction speed for the additional tolerance of noise. The first order FPM reconstruction algorithm, which was much faster than the second order algorithm, performed worse than the other two in both cases. This is in accordance with previously reported results, which suggested the second order algorithm is much more robust than the first order alternating projection method [9].

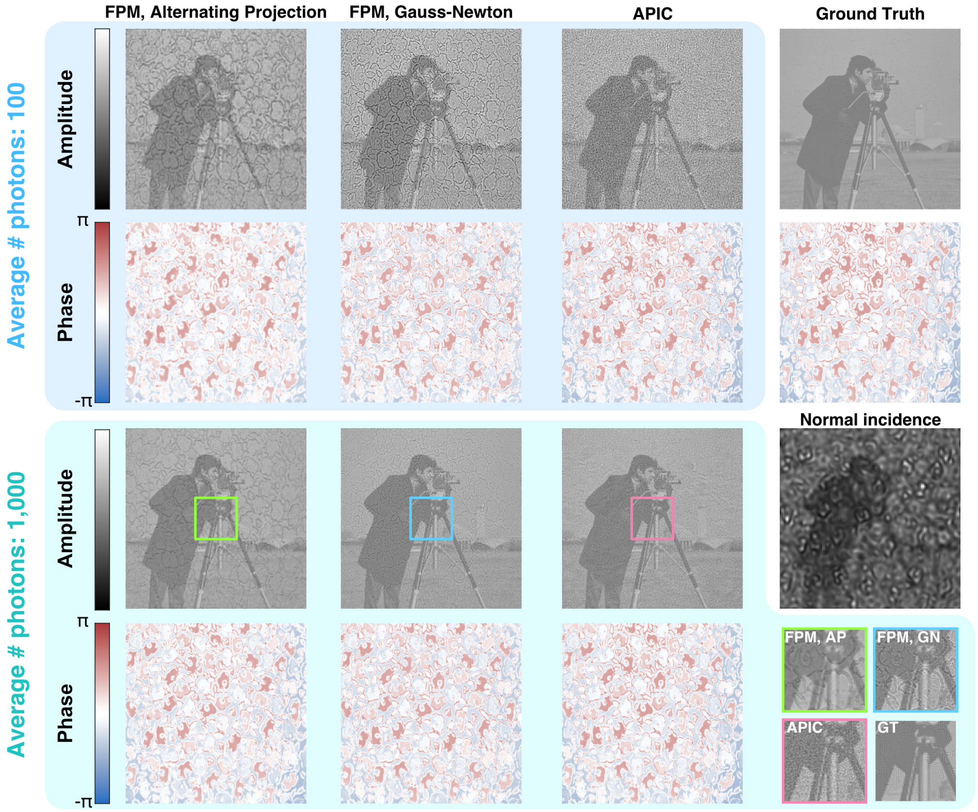

Figure 2.23: Reconstruction of a weak absorption target using FPM and APIC. The numbers shown on the left are the average number of photons under the matching angle. The zoomed version of images inside the highlighted boxes, along with the zoomed ground truth, are shown on the lower right of this figure. The image under normal incidence was simulated with a noise-free imaging system. For a complex sample whose amplitude and phase are different in morphology, cross-talk between the two is apparent in FPM's reconstruction. APIC, however, does not suffer from such severe cross-talk.

In our second simulation, we chose two different patterns for the amplitude and phase of the complex object. This complex object was designed to have a weak amplitude variation while preserving a relatively strong phase variation based on the common property for unstained biological samples. We see that there were severe cross-talks between phase and amplitude in both FPM algorithms using low SNR measurements. When SNR increased, such cross-talk became less prominent in FPM. Nonetheless, structures of the phases remained visible in those cases. APIC, on the contrary, did not suffer from such cross-talk. Although the reconstruction of APIC with low SNR was noisy, it maintained the features of the ground truth amplitude and phase and showed almost no cross-talk between the two. Another thing we found in our simulation is that the reconstructed phase of APIC was closer to the ground truth. While the structural features of the reconstructed phase of FPM were similar to the ground truth, they quantitatively differed from each other. For the low SNR dataset, the range of the reconstructed phase of FPM appeared to be compressed. This indicates that if the reconstruction artifact is undesired and the fidelity is of concern, APIC is definitely a preferable choice when imaging most unstained biological samples, even under low SNR conditions.

Number of NA-matching measurements required in APIC

In this section, we simulated different numbers of NA-matching measurements to see the required measurements for APIC to accurately reconstruct the imaging system's aberration. In our simulation, we only introduced aberration to the imaging system and all other parameters were assumed to be ideal. We assumed the illumination orientations were azimuthally uniformly distributed, which means their corresponding LEDs were uniformly distributed along a certain ring. The reconstructed aberrations are shown in Fig. 2.24.

When using 4 images, APIC did not obtain a good aberration estimate. However, we can see that when there were 6 NA-matching measurements, APIC successfully reconstructed mild to moderately high aberrations (their phase standard deviation is below 0.8π). However, residual aberration exists in the reconstructed aberration under severe aberrations (phase standard deviation exceeds 1.5π), as depicted in the corresponding error map in Fig. 2.24. When the number of NA-matching measurements increased to 8, we see that there are no residual aberrations in APIC's reconstructions regardless of the severeness of the aberration. In general, 6 to 8 NA-matching measurements would be sufficient for generating an accurate aberration estimation in most cases.

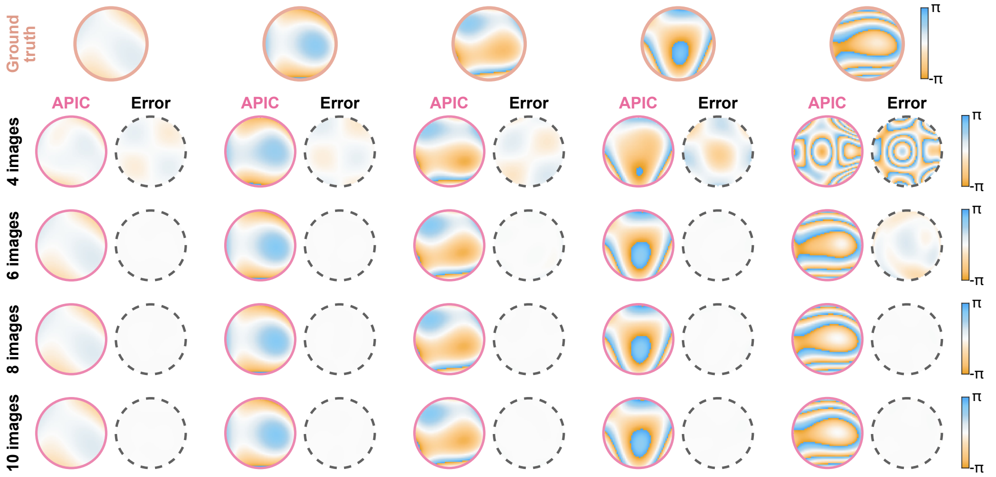

Figure 2.24: Recovered aberration and the number of NA-matching measurements. The ground truths in the simulation are shown on top of each group. The number of NA-matching measurements used in the simulation are dented on the far left size of the image. In addition to the reconstructed aberration, we also included the error map between this reconstructed aberration and the ground truth. This error map is placed next to the APIC's reconstruction and is highlighted by a black dashed circle.

Inaccurate illumination angle estimates

We simulated errors in the angle calibration to see how APIC performs under different levels of calibration error. In our simulation, random uniformly distributed estimation error was introduced to the actual illumination angles. For different levels of severeness of the calibration errors, the maximal amount of error in our simulation was proportional to the maximal acceptance angle of the imaging system (that is, the error is proportional to k_{NA} of the system). As the illumination angle was converted to the spatial frequency vector in practice, this maximum error is denoted by the ratio of the error to the k-vector of the maximal acceptance angle. In our simulation, we assumed the estimate of the NA-matching angle was no larger than the acceptance angle of the imaging system.

From the simulation results, we can see that APIC performance was well-maintained until the error reached 9%. Beyond that, the reconstructions showed obvious artifacts. This indicates that the alignment requirement of our proposed APIC is fairly relaxed. In other words, APIC is tolerant to the LED position errors to a good extent.

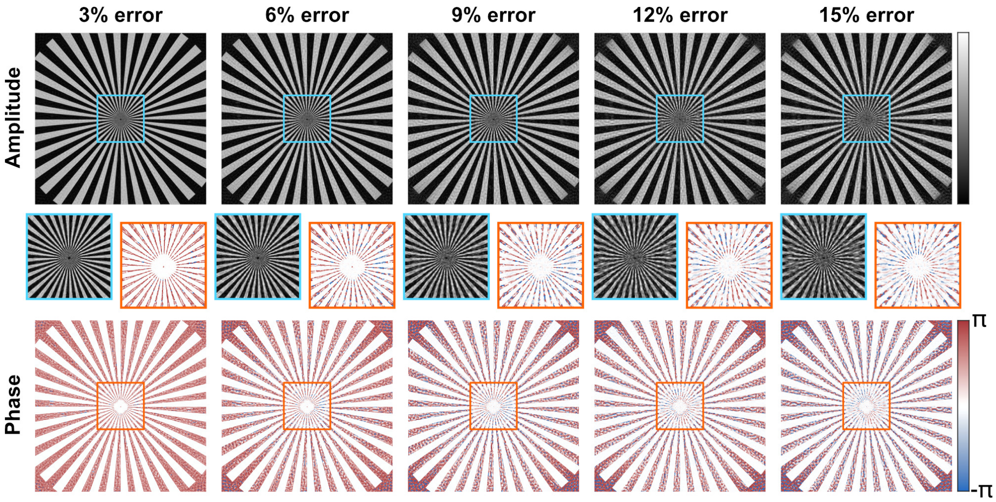

Figure 2.25: APIC reconstruction with different levels of angle error. In this simulation, we added randomly generated angle errors to the actual illumination angles and fed this inaccurate angle estimate to APIC when performing the reconstruction. The numbers on top of each group are the maximal possible error in our simulation. If the maximal possible error is denoted by γ, the simulated angle errors for both the horizontal and vertical direction were uniformly sampled in $[-\gamma, \gamma]$. For each NA-matching measurement, we forced the corresponding angle estimate to be smaller than the acceptance angle of the system by tuning the sign of this randomly generated error.

Chapter 3

ULTRASOUND AIDED IMAGING THROUGH SCATTERING: AN APPLICATION

This chapter is adapted from the manuscript, Liu Y., Cao, R., Xu, J., Ruan, H., and Yang, C. “Imaging through highly scattering human skulls with ultrasound-modulated optical tomography,” *Optics Letters 45, 2973–2976 (2020).*

In this chapter, we start to deal with scattering. We showed in Fig. 1.13 that the original focus will be destroyed in the presence of scattering medium. Thus, it is of great difficulty to reconstruct the image we want without extra designs in the imaging system. Furthermore, in the case of scattering, the CTF is no longer a good tool to model the image process because points at different locations are typically imaged to completely different patterns when interacting with a scattering medium. As a result, some other types of methods should be used in counteracting scattering and obtaining images of the sample of interest.

To see what are the possible ways to achieve such goal, we dig into the mechanism of how scattering destroys a conventional focus and try to find some ways from there. As the ramdom phasor sum is the cause of the speckle patterns (Fig. 1.12), we want to minimize this random phase modulation introduced by the scattering medium to “reduce” scattering. Given a certain sample, the difference in the refractive index is a slow varying function in the visible light regime. Moreover, it turns out that the absolute path length difference is approximately the same for all visible light for most samples. For light with a longer wavelength, the phase change induced by the same path length difference is thus much smaller compared with that of a much shorter wavelength: Let the path length difference be Δd and the wavelength be λ, the phase change $\Delta\varphi$ induced by the path length difference is given by

$$\Delta\varphi = 2\pi\frac{\Delta d}{\lambda}. \tag{3.1}$$

Given a fixed path length difference, a larger wavelength then corresponds to a smaller phase change. This intuition applies to other waves in general. Thus, to mitigate the negative effect of scattering, one straightforward idea is to use waves of much longer wavelengths compared with light. When the wavelength is much larger than the path length difference, the induced phase change becomes negligible and scattering effect can be omitted in analyzing this wave. Based on this idea, we will explore the usage of sound in optical imaging in the following section.

3.1 Background on offaxis hologram and UOT

As the absorption properties of a sample in the visible light regime and the ultrasonic regime are different, direct usages of the ultrasound will not carry information of optical properties. Thus, a way to use ultrasound in an optical imaging system is desired. There are two prominent ways of introducing ultrasound for visible light: light can induce sound waves and the ultrasound is able to modulate light's frequency. When a material, say tissue, absorbs a pulsed light, the material is heated up first and then cooled down to normal temperature. This thermal change results in the expansion and contraction of the material and ultimately produces an ultrasonic wave. This light induced sound process is called the photoacoustic effect. We note that there are some other mechanisms in the generating of sound and interesting readers can refer to literature for more information. Previous literature also show that ultrasound could induce a frequency change in light: when interacting with the ultrasound, the frequency of the light changes according to the ultrasonic frequency. By utilizing this frequency change, we can filter out the light that does not interact with ultrasound.

For photoacoustic effect, the induced sound wave should be detected by sound transducers, which convert the sound signal to the electric signal that can then be recorded (in other words, these are "sound detectors"). For ultrasound induced frequency change, we can still use a light detector as the ultrasound modulates the light and the information is carried by the light. Two representative methods using photoacoustic and ultrasound modulation in optical imaging are photoacoutstic tomography (PAT) [23] and ultrasound-modulated optical tomography (UOT) [24], respectively. In this section, we focus on the latter as it is easier to integrate in most optical imaging system.

As light detectors only capture the intensity of the light field and its phase and frequency information cannot be directly measured, we need to find a way to extract the frequency shifted light. Here, we demonstrate one popular complex field detection method: the offaxis hologram. We explained in Section 2.1 that a tilted reference beam is used in generating the hologram, which enables complex field reconstruction. Here, its detailed mechanism will be discussed. We know from Eq. 1.17 that when two different fields interfere, the recorded intensity varies based on the phase differences of two fields. This can be utilized to extract the complex field. We will include the time-varying part in the following section as we are going to deal with multiple frequencies. Without loss of generality, we assume the reference field to be a plane wave with phase offset zero and let its tilt be **k**, we can write the reference field $E_r(\mathbf{x}, t)$ as

$$E_r(\mathbf{x}, t) = A_r e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)}. \tag{3.2}$$

Let the ultrasound modulated field be $E_m(\mathbf{x}, t)$ given by

$$E_m(\mathbf{x}, t) = A_m(\mathbf{x}) e^{i(\phi_{\mathbf{x}} - \omega t)}, \tag{3.3}$$

and let the light that does not interact with ultrasound be $E_b(\mathbf{x}, t)$ with a frequency ω' that is shifted from ω ($\omega' \neq \omega$)

$$E_b(\mathbf{x}, t) = A_b(\mathbf{x}) e^{i(\phi'_{\mathbf{x}} - \omega' t)}. \tag{3.4}$$

We first look at the time-varing intensity of the interference of $E_m(\mathbf{x}, t)$ and $E_b(\mathbf{x}, t)$, which is given by

$$\begin{aligned} I_{mb}(\mathbf{x}, t) &= \left|E_m(\mathbf{x}, t) + E_b(\mathbf{x}, t)\right|^2 = \left|A_m(\mathbf{x}) e^{i(\phi_{\mathbf{x}} - \omega t)} + A_b(\mathbf{x}) e^{i(\phi'_{\mathbf{x}} - \omega' t)}\right|^2 \\ &= A_m^2(\mathbf{x}) + A_b^2(\mathbf{x}) + 2A_m(\mathbf{x})A_b(\mathbf{x}) \cos\left(\phi_{\mathbf{x}} - \phi'_{\mathbf{x}} - \omega t + \omega' t\right). \end{aligned} \tag{3.5}$$

As the ultrasound induced frequency shift is also much bigger than the response of the detector, the detector captures the time-averaged intensity

$$\begin{aligned} \mathbb{E}_t\left[I_{mb}(\mathbf{x})\right] &= A_m^2(\mathbf{x}) + A_b^2(\mathbf{x}) + 2A_m(\mathbf{x})A_b(\mathbf{x})\mathbb{E}_t\left[\cos\left(\phi_{\mathbf{x}} - \phi'_{\mathbf{x}} - \omega t + \omega' t\right)\right] \\ &= A_m^2(\mathbf{x}) + A_b^2(\mathbf{x}). \end{aligned} \tag{3.6}$$

where $\mathbb{E}(\cdot)$ denotes the expectation and its subscript t denotes the expectation is taken with respect to time. Thus, the detected intensity is just the sum of the intensity of the two fields. Note that this applies to the reference field as well because the modulated field exhibits the most general form of a complex light field. This suggests two fields with different frequencies “do not” interfere with each other when a detector is slower than their frequency difference.

Following the same calculation, the interfered field of the reference and the modulated field is given by

$$I_{rm}(\mathbf{x}, t) = A_r^2 + A_m^2(\mathbf{x}) + 2A_r A_m(\mathbf{x}) \cos\left(\mathbf{k} \cdot \mathbf{x} - \phi_{\mathbf{x}}\right). \tag{3.7}$$

The cross part can be rewritten as

$$2A_r A_m(\mathbf{x}) \cos\left(\mathbf{k} \cdot \mathbf{x} - \phi_{\mathbf{x}}\right) = A_r A_m(\mathbf{x}) e^{i(\mathbf{k}\cdot\mathbf{x}-\phi_{\mathbf{x}})} + A_r A_m(\mathbf{x}) e^{-i(\mathbf{k}\cdot\mathbf{x}-\phi_{\mathbf{x}})}. \tag{3.8}$$

We can then apply the shift property of Fourier transform (Table 1.1) to conclude that this cross term shifts the (scaled) Fourier transform of $A_m(\mathbf{x})e^{i\phi_x}$ by $\pm\mathbf{k}$ in the spatial frequency domain. If the shift is larger than the maximal spatial frequency of both $A_m^2(\mathbf{x})$ and $A_b^2(\mathbf{x})$, the cross term can be extracted from the intensity measurement, as shown in Fig. 3.1. This is the mathematical intuition behind the offaxis hologram. As the sample is typically imaged using an imaging system, the sampled light field (including light does not interact with the ultrasound and light with ultrasound modulation) has a maximal spatial frequency that is limited by the NA of this imaging system. Thus, we can build an imaging system that satisfies the aforementioned requirement and reconstruct the complex field from the intensity measurement. Figure 3.1 includes a typical intensity measurement and its Fourier transform in offaxis holography.

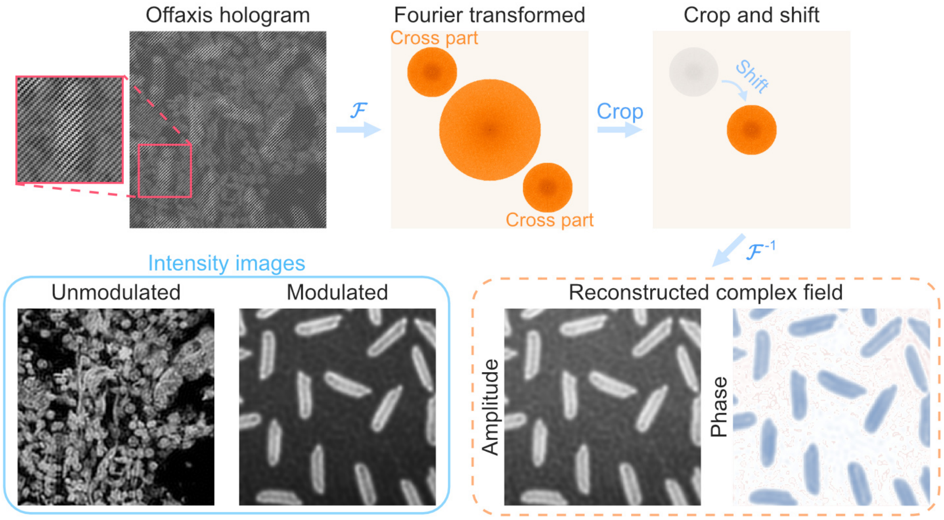

Figure 3.1: Reconstruction in offaxis hologram. When a proper tilting angle is chosen for the reference field, the two cross terms are away from the central circle in the spatial frequency domain. As we showed that this cross term is a scaled version of the ultrasound modulated field, it can be extracted to reconstruct the modulated light field. The intensity images correspond to the unmodulated light and the ultrasound modulated light are shown in the bottom left of this figure.

Building upon this method, the light field whose frequency is shifted by the ultrasound can be extracted from an intensity measurement. This technique can be used in detecting the frequency-shifted light in UOT. In UOT, the ultrasound is focused on the sample and raster scanned over the sample. As only light interacting with the ultrasound is extracted, the energy of the extracted light reflects how absorptive the sample at the ultrasound focus is. Thus, by scanning the ultrasound focus across the sample, the sample's optical absorption map can be obtained. Using UOT, optical properties can be extracted with a resolution that is determined by the ultrasound focus.

UOT has been utilized in many scattering related applications because sound is inherent less sensitive to scattering compared with light. However, application of UOT in ultrahigh scattering regime remains largely unexplored. Here we investigate the possibility of imaging through highly scattering human skulls with UOT. We used a single-shot off-axis holography method [25] to detect the UOT signal, which to our knowledge has not been demonstrated in the UOT field. The large pixel count of a camera enables parallel detection of multiple speckle grains within a single shot, which is crucial for this type of applications where the signal-to-noise ratio is low and the speckle decorrelates rapidly due to physiological motions such as blood flow [26–28]. In previous work [29], multiple frames were recorded to reconstruct the UOT signal, making the method vulnerable to the rapid speckle decorrelation in living tissue. Single-frame UOT measurements have previously been reported [30–32]. However, some method requires special lock-in cameras in which each pixel is an analog lock-in detector [31–33]. Although off-axis holography has been used in heterodyne holography based UOT, multiple frames were required to filter out the local oscillator beam [34].

3.2 UOT results

In our experiments, we first evaluated the extent by which a human skull distorts the ultrasonic field. A hydrophone (HNR-0500, ONDA) was used to measure the ultrasonic field distribution on the focal plane of a spherically focused single-element ultrasonic transducer (A303S, Olympus; central frequency = 1 MHz, focal length = 15.2 mm, element diameter = 12.7 mm) with and without the presence of a human skull (3-5 mm thick, human parietal bone, SHN-46, Skulls Unlimited International Inc, transport mean free path is approximately 1 mm). Figure 3.2a shows a photo of the setup and Figure 3.2b shows the normalized ultrasonic field distribution. The full width at half maximum (FWHM) focal spot size was measured to be 3 mm when the skull was absent. When the skull was present between the transducer and the hydrophone, depending on the location and orientation of the ultrasonic transducer relative to the skull, the FWHM focal spot size varied from 3 mm to 6 mm, and the ultrasonic pressure at the focus was attenuated to 10%-20%. Although distorted and broadened, we observed that a focus was still achievable through the human skull for 1 MHz ultrasound — a result that is consistent with previous literature findings [30].

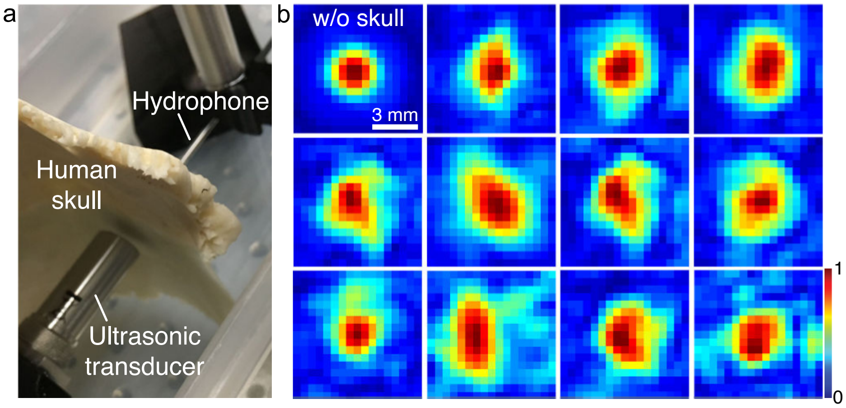

Figure 3.2: Spatial varying ultrasound foci. We measured the ultrasonic field distribution on the focal plane of a spherically focused transducer, with and without the presence of a human skull. **a**, Photo of the experimental setup. **b**, The measured ultrasonic field distribution (normalized) at different lateral locations of the ultrasonic transducer relative to the fixed human skull. Note that the bone thickness is not uniform across the skull.

Next, we built a camera based UOT system (schematically shown in Fig. 3.3) to demonstrate the feasibility of imaging an absorptive object buried in between two pieces of highly scattering human skull. The output of a continuous-wave laser (671 nm, MSL-FN-671-S, CNI Optoelectronics Tech Co.; ~35 mW on the sample) passed through an optical isolator, a variable attenuator composed of a half-wave plate (HWP1) and a polarizing beamsplitter (PBS1), before it was split into a reference beam (R) and a sample beam (S) by a polarizing beamsplitter (PBS2). After passing through a neutral density filter (ND), the reference beam was expanded to a diameter of 1 inch by two lenses (L1 and L2), and reflected by a 90:10 (T:R) non-polarizing beamsplitter (BS) before it illuminated a camera sensor. The sample beam passed through a half-wave plate (HWP3), two acousto-optic modulators (AOM1 and AOM2, which shifted the frequency of the sample beam by 50 MHz and -49 MHz sequentially), and a beam expander, before it illuminated the human skulls. An ultrasonic transducer focused 1 MHz ultrasound through the skull and the focal pressure is ~0.34 MPa. A portion of the light passing through the ultrasonic focus was tagged by the ultrasound and its frequency was shifted to the same frequency as the reference beam. The ultrasonically tagged light and untagged light passed through an absorptive object (a strip of 2 mm wide and 3 cm long black tape, transmittance < 0.1%, attached on the skull), a second human skull (1.5 cm away from the first skull), a 4f system (with an iris in the pupil plane to adjust the speckle size), a polarizer, and the BS before they interfered with the reference beam (45° with respect to the x and z axes) and detected by a camera (pco.edge 5.5, PCO-TECH; global shutter, 2560 × 2160 pixels, 6.5 µm pixel size). Because the ultrasonically tagged light had the same frequency as that of the reference beam, their interference pattern was stable on the camera. In contrast, the interference pattern formed by the untagged light and the reference beam was a 1 MHz beat, and thus was averaged out during the camera exposure time (5 ms).

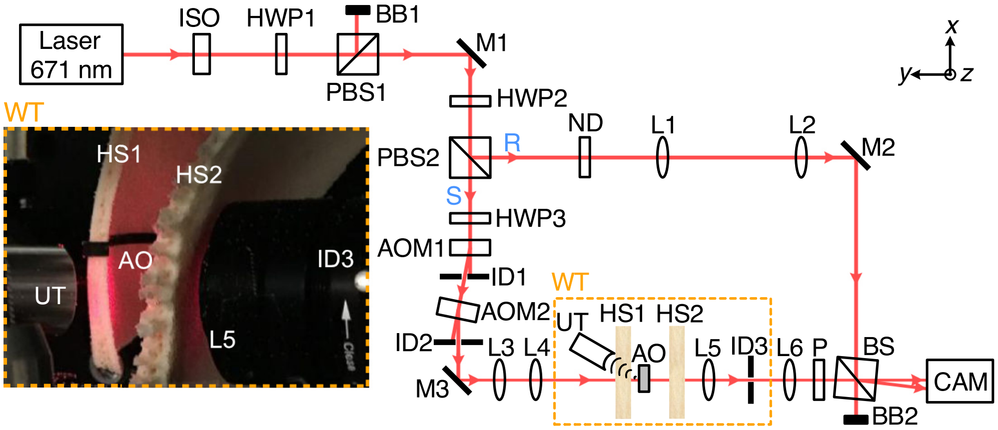

Figure 3.3: Schematic of the UOT setup. Abbreviations: AO, absorptive object; AOM, acousto-optic modulator; BB, beam block; BS, non-polarizing beamsplitter (90% transmission, 10% reflection); CAM, camera; HWP, half-wave plate; HS, human skull; ID, iris diaphragm; ISO, isolator; L, lens; M, mirror; ND, neutral density filter; P, polarizer; PBS, polarizing beamsplitter; R, reference beam; S, sample beam; UT, ultrasonic transducer; WT, water tank. The components in the water tank are enclosed in a dashed box, and a photo is shown.

To obtain an image of the absorptive object, we scanned the ultrasonic focus with respect to the object along the x direction (see the labeling of axes in Fig. 3.3) and measured the UOT signal at each scanning position. By plotting the normalized UOT signal as a function of the ultrasonic focus position, we obtained an image of the object, which is shown in the blue solid curve in Fig. 3.4. The dip of the curve at ~6 mm position manifests the object, because the object absorbed some of the light. When we removed the object and repeated the measurement (see the red dotted curve in Fig. 3.4), we did not observe a similar dip and the curve is relatively flat, showing that the image contrast in the blue solid curve is mainly caused by the object, rather than the skull inhomogeneity.

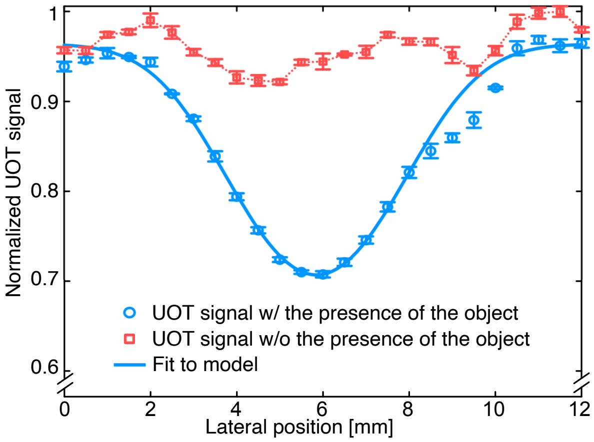

Figure 3.4: UOT signal as a function of the ultrasonic focus position. The blue circles denote the signal we acquired with an absorptive object (a black slit in this case), and the red squares denote the signal without the object. The blue solid curve denotes the fitting to the experimental data with the forward model shown in Eq. 3.10. Each data point is an average of four measurements and the associated error bar represents the standard deviation.

In UOT, the point scanning strategy is used to form the final absorption map. As such, the images acquired at different times do not interfere with each other. Thus, this system obeys a slightly different model compared with the coherent imaging model in Eqs. 1.11 and 1.12. In incoherent case, each point on the sample is ought to be mapped to another "point" and the intensity is directly summed up. This is to say, the system acts on the intensity of the sample instead of the complex field. Based on a similar discussion in establishing the CTF, we have the following forward model for incoherent imaging

$$I(\mathbf{x}) = \mathcal{F}^{-1}\left[\hat{O}(\mathbf{k})H_i(\mathbf{k})\right](\mathbf{x}), \tag{3.9}$$

where I is the acquired intensity, $\mathbf{x}$ describes the scanning location, $\mathbf{k}$ is the spatial frequency vector, $\hat{O}(\mathbf{k})$ denotes the Fourier transform of the sample's absorption map $O(\mathbf{x})$ and $H_i(\mathbf{k})$ is a function that is similar to CTF but instead used in incoherent imaging. We call function $H_i(\mathbf{k})$ the optical transfer function and its inverse Fourier transform the point transfer function (PSF). In terms of the PSF, the imaging system can be modeled by

$$I(\mathbf{x}) = O(\mathbf{x}) * h(\mathbf{x}) \tag{3.10}$$

where $*$ denotes convolution, $O(\mathbf{x})$ represents the object and can be modeled by a rectangular function with a width of 2 mm in our experiment, and $h := \mathcal{F}^{-1}(H_i)$ is the PSF of the imaging system. We note that the PSF in our experiment can be modeled by a Gaussian function with a FWHM size related to the system resolution determined by the ultrasonic frequency, numerical aperture, and skull distortion. By fitting the experimental data with the imaging model (see the blue solid curve in Fig. 3.4), the FWHM system resolution is found to be 4.1 mm, which is within the normal range considering the skull induced distortion of the PSF (see Fig. 3.2). Misalignment of the ultrasonic focus position with respect to the object along the y direction also degrades the measured lateral resolution.

3.3 Discussion on UOT for imaging through scattering medium

In this work, we use a hybrid technology combining light and sound to image through highly scattering human skulls. Our method images optical contrast at depths, with ultrasound determined spatial resolution. Therefore, it can achieve a higher spatial resolution than that of diffuse optical tomography [26]. Currently, the resolution along the acoustic axis direction is poor, because we used a continuous-wave (cw) laser (and thus a long burst of ultrasound) in this experiment. The axial resolution can be improved by using a pulsed laser and a single cycle of ultrasonic pulse [35]. This can also improve the lateral resolution and image contrast because the tagging volume is much bigger in the cw case and the beam widths at out-of-focus planes are larger than that at the focal plane. The dip in Fig. 3.4 is caused by both the absorption of tagged light and the absorption of the incident untagged light.

In our experiment, we used a single element transducer to generate an ultrasonic focus through the human skull. Better focusing quality (e.g., smaller focal spot and higher focus to background pressure ratio) can be achieved with a high numerical aperture transducer array and acoustic wavefront shaping to correct for the skull induced aberration (i.e., adjusting the delay of each element to let the fields associated with different elements constructively interfere at the focus) [36]. Due to reciprocity of acoustic waves, PAT can in theory achieve the same spatial resolution through the skull as UOT. It would require PAT methods to correctly find the delay of each transducer element. One way to find such delays in PAT would be to use a computational model which incorporates the skull morphology and composition information obtained from x-ray computed tomography (XCT) [37]. However, the effectiveness of this method relies on precise registration of the two imaging modalities involved (i.e., XCT and PAT) and the accuracy of the skull model.

While this method works well for monkey skulls [37], its effectiveness for adult human skulls remains to be demonstrated experimentally. In comparison, it may potentially be easier to find such delays experimentally in UOT experiments. For example, one can adjust the delay of each element until the nonlinear UOT signal or the temperature rise at a target location monitored by magnetic resonance imaging is maximized [38]. In such UOT experiments, we would also avoid the need to perform registration with XCT — another source of potential error in PAT.

We demonstrate the feasibility of imaging an object through human skulls using a transmission-mode system. Because human frontal bones and parietal bones are highly curved, our experiment mimics the case where we illuminate from one location of the bone and detect the transmitted light at an offset location from the illumination site. For other non-curved bones, a reflection-mode system should be used.

The speckle size on the camera and the angle of the reference beam should be controlled so that the sideband in the Fourier space does not overlap with the zeroth order, while the area of the sideband should be maximized to capture more speckles to increase the signal. The speckle size on the camera was 5.6 pixels wide in our experiment, and it can be further reduced to 4 pixels wide to increase the signal [39]. In addition, rather than using a circular iris, a rectangular iris can be employed to maximally use the Fourier space (camera pixel count) in off-axis holography [34].

The off-axis holography based UOT detection method has a high sensitivity. Because interferometry is used to boost the signal above the detector noise, the detection is shot-noise limited. The expression of the SNR in offaxis based UOT can be derived and written as (See Appendix A)

$$\mathrm{SNR}_{\mathrm{cam}} = \frac{\mathrm{UOT}_{\mathrm{signal}}}{\sqrt{\mathrm{Var}(\mathrm{UOT}_{\mathrm{signal}})}} = \frac{\sqrt{N_P}\bar{N}_T}{\sqrt{4\bar{N}_T + 1}}, \tag{3.11}$$

where $\bar{N}_T$ is the average number of ultrasonically tagged photons per pixel and N_p is the pixel count of the camera. In deriving Eq. 3.11, we assumed that the number of reference beam photons per pixel equals the average number of untagged photons per pixel, and both numbers are much larger than $\bar{N}_T$, which is commonly achieved in experiments. We also assumed a quantum efficiency of 1 and a rectangular iris is employed. Assuming $N_p = 106$, Equation 3.11 shows that the SNR is above one as long as $\bar{N}_T > 1/\sqrt{N_p} = 10^{-3}$.

Chapter 4

NON-LINE-OF-SIGHT IMAGING USING ACTIVE FOCUSING

This chapter is adapted from the manuscript, Cao, R., de Goumoens, F., Blochet, B., Xu, J. and Yang, C. "High-resolution non-line-of-sight imaging employing active focusing," *Nature Photonics 16, 462–468 (2022).*

In the previous chapter, we investigated the use of ultrasound in optical imaging. We can see the resolution of this method is primarily determined by the sound wave and is poorer than most optical methods. In this chapter, we will instead try to directly focus the light in the presence of the scattering medium. We demonstrate our method on one interesting problem, namely the non-line-of-sight problem. The setting of this problem and previous works on addressing this problem are presented in the following section.

4.1 Introduction of the non-line-of-sight imaging problem

In this section, we introduce the non-line-of-sight (NLOS) imaging problem, which aims to solve the scattering related image reconstruction problem [40]. We call places where we can see without any obscure the line-of-sight range. The non-line-of-sight, on contrary, refers to places where we cannot directly see. A typical NLOS imaging scenario can be described as follows: The target object is hidden from an observer's direct line-of-sight view and only light from the non-line-of-sight path is detected (see Fig. 4.1). Typically, the light is passively scattered by the wall and then being detected. The scattering of the wall prevents the observer from collecting a clear image of the object. Nevertheless, it has been shown that it is possible to image the object with novel optical approaches, such as speckle correlations [41, 42], time of flight (ToF) imaging [43–45], phasor field [46–48], and other computational imaging methods [49]. Here, we will give a short introduction on the time-of-flight method, as it is one of the most popular method for addressing the non-line-of-sight problem. Please refer to the corresponding literature for the principle and implementation of the several other techniques mentioned here.

In the ToF method, a fast time-resolved detector is used to record the traveling time of light. This traveling time can then be converted to the distance. The core idea is to use the differences in the distance (in other words, path length differences) to locate different points in a three-dimensional space, as shown in Fig. 4.1. This principle can be seen from the following: When the laser is oriented to the wall, the distance between the light spot on the wall and the laser is fully determined (and controllable). Let the detector be imaged on a particular position on the wall, the distance between this detector and the point that reroute the light into the detector is also known. By extracting the distance light travels, we know $R_0 + R_1$. We can then construct an ellipsoid using this known distance. For an unknown sample, each point on the sample induces a different light path, which ultimately results in a different photon arrival time. Thus, by changing the geometry of the system, many more of those ellipsoids can be determined and then be used for the image reconstruction.

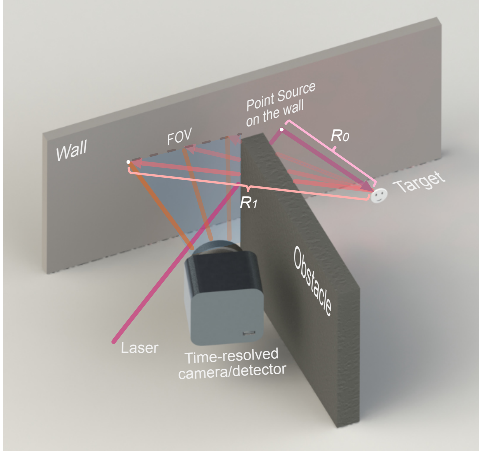

Figure 4.1: Concept of a time-of-flight system. In ToF system, a short pulse of laser is shined onto the wall. The laser is scattered by the wall, bounced back by the target (say, by a certain point), and then scattered by the wall for a second time. The fast, time-resolved detector then collects the light and record the time of flight on the light pulse. Use this information, we know the sum of distance R_0 and R_1. Thus, with a single pair, we find an ellipsoid that contains the point on that target that scatters the light. Using multiple pairs to generate multiple different ellipsoids, we can uniquely identify the location of that point.

Although we can use ToF to image the hidden target in NLOS, there are several issues with exiting methods. One big problem with existing NLOS methods without using prior information of the wall is that they do not provide diffraction limited resolution. These NLOS imaging methods that are implemented at a practical wall-object distance (> 0.1 m) generally produce modest distance-to-resolution ratios (approximately the inverse of the angular resolution in radian, the distance here is the wall-to-object distance) — a factor of approximately 100 is the best number being reported so far [44–47]. Moreover, these methods do not directly counteract scattering. Instead, they extract useful information from complex signals that usually come with a big background.

These limitations motivate us to investigate ways to directly focus light on the hidden object in the NLOS setting.

4.2 The wavefront shaping technique

Before we introduce our active focusing based NLOS imaging method, we will first introduce the wavefront shaping technique, which aims to undo the random phase modulation induced by a scattering sample [50–63].

We know from Chapter 1 that the original incident wavefront is destroyed after interacting with a scattering medium (Figs. 1.4 and 4.2a). To generate a focus in the presence of scattering, we want to obtain a focusing wavefront after the light interacts with the scattering medium. This is what wavefront shaping can do: it reshapes the wavefront of an arbitrary optical system, including one with scattering media.

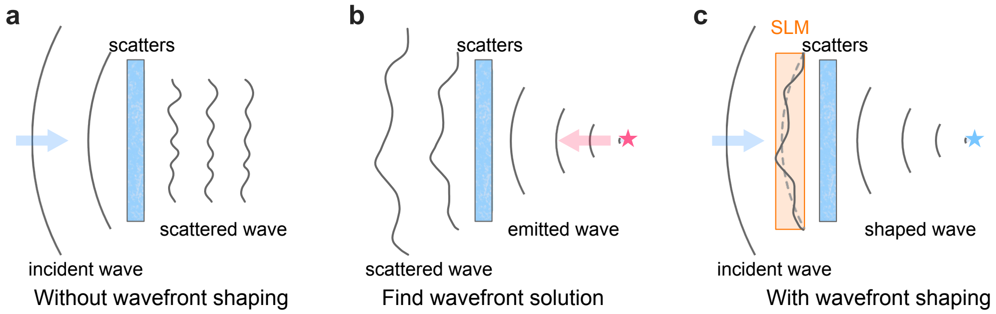

Figure 4.2: Principle of wavefront shaping. **a**, Insert a scattering medium in a focusing beam. Due to the presence of the scattering medium, the original spherical wavefront is destroyed. **b** The scattered wavefront emitted by a point source. Similar to what happens to a focusing beam, the emitted wavefront from a point source is modulated by the scattering medium and the modulated wavefront does not possess a regular shape. **c**, Employ wavefront shaping to reproduce the spherical wavefront after interacting with the scattering medium. If we put the scattered wavefront of a point source shown in **b** back to the scattering medium, the light should in principle trace back to a point. Thus, if there is a way to find the scattered wavefront of a point source, we can shape the wavefront accordingly so that the reshaped wavefront is transformed into a focusing wave when interacting with the scattering medium. SLM: spatial light modulator.

Wavefront shaping is developed based on optical reciprocity: if a light ray passes through a fixed linear, non-magnetic optical system and shows up at point Y, then it will travel back along the same trajectory if we reverse the light at point Y. This principle also applies to the wavefront of light. Let us first assume we know the scattered wavefront of the light emitted by a point source and modulated by a scattering medium (Fig. 4.2b). Then, if we reverse the scattered light, the reserved light field should be transformed back to the original spherical wave (with the opposite propagating direction) when it interacts with the same scattering medium. Thus, if we use a spatial light modulator (SLM) to reshape the initial incident wave in Fig. 4.2a to the scattered wavefront from a point source, the engineering wavefront should now produce a focus when interacting with the scattering medium, as shown in Fig. 4.2c.

To find the "best" wavefront, we can measure the scattering properties of the random medium. If we know the phase modulation the scattering medium imposes on the entire space, we can use that to infer the desired wavefront. In wavefront shaping field, we use "transmission matrix" to characterize the properties of a scattering medium and several transmission matrix based approaches demonstrated the feasibility of focus generation through a scattering medium [53–55].

Another approach is to directly find the scattered wavefront of the point source shown in Fig. 4.2b, which is one of the most fundamental thing in guidestar based wavefront shaping [58–63]. In the past few decades, several methods in generating this desired point source (which is called "guidestar" in wavefront shaping) were proposed. For a single fluorescence guidestar, we can find the wavefront that leads to a focused spot by maximizing the fluorescence signal. We can also use the ultrasound as the guidestar: light interacting with the ultrasound focus will be frequency-shifted and the frequency modulated wavefront should trace back to the ultrasonic focus. Other mechanisms in extracting the desired focusing wavefront are also explored and interesting readers can refer to literature for their detailed implementation [59].

We note that there are prior wavefront shaping based NLOS imaging works that involve measuring the wall's wavefront scattering characteristics prior to inserting a hidden object [64, 65]. One work [64] uses a point source as the initial guidestar, while the other work [65] uses a camera as a guidestar proxy. Spatial light modulators are used to correct the scattered wavefront and a linear phase ramp is applied to raster-scan the generated focus [65–67]. These works assume the scattering properties are known, which is not a practical assumption in a general NLOS imaging problem. Thus, we need to develop a method that is applicable in a general NLOS setting.

4.3 Introduction and principle of UNCOVER

With the benefit of wavefront shaping, we can in principle use wavefront shaping to solve the aforementioned NLOS imaging issues. However, the straight translation of wavefront shaping technology for NLOS imaging is infeasible in a general NLOS setting as we have no access to hidden target's side and thus cannot implement a guidestar on the target. Thus, ways to "find" a guidestar in a general NLOS setting are needed.

Here, we introduce a new wavefront shaping method, termed Unseen Non-line-of-sight Casted Optical aperture Visibility Enhanced Return (UNCOVER) focusing. UNCOVER uses the hidden object itself as the guidestar and generates a near-diffraction-limited focus that is substantially smaller than the object, which can be raster-scanned to image the object (with its actual 3D coordinates remaining unknown). We note that there are prior wavefront shaping works that can produce a focus that is smaller than the guidestar [61–63], UNCOVER differs from these methods in that it does not require direct modulations (such as ultrasonic modulation) on the object. Thus, it is well suited for the NLOS settings.

In UNCOVER, we measured the phase variation induced by the wall on the reflected light, and engineer a wavefront that focuses at the target when reflecting off the wall. Unlike general wavefront shaping methods, UNCOVER uses the fact that the wall is a 2D thin scattering object. By imaging the wavefront onto the wall, the angular aperture of the illumination from the wall to the object can be controlled in the wavefront shaping process. Through this process, UNCOVER is able to generate a tight optical focus (smaller than the hidden object) at the object.

An illustration of the UNCOVER system is shown in Fig. 4.3. The system is optically configured so that the spatial light modulator (SLM), the adjustable input pupil (implemented with a digital micro-mirror device, DMD, in our experiment) and the wall are an image-forming conjugate plane set. The light transmitted through the SLM is imaged onto the wall and its diffuse reflection illuminates the object. The return light from the object scatters off the wall once again. This return light is not wavefront-shaped and is diffusely collected by a detector (see Fig. 4.3) — we can reasonably assume that the detected signal is proportional to the total light power reflected by the object.

In UNCOVER, our goal is to use the SLM and DMD to project an engineered wavefront that can counter the spatial varying phase shift associated with reflection from the wall towards the object. The reflection off the wall will render a tight

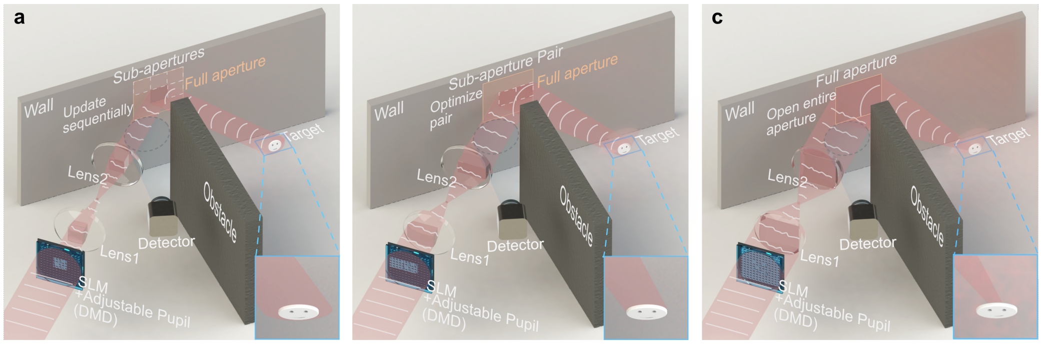

Figure 4.3: Concept of UNCOVER's system setup. A spatial light modulator (SLM) is imaged onto the wall and modulates the transmitted wavefront. A adjustable pupil (implemented with a digital micromirror device, DMD) is used to control the illuminated area (aperture) on the wall. **a**, UNCOVER firstly optimizes the wavefront transmitted by a small pupil (corresponds to a sub-aperture on the wall) to focus light onto the target. Optimal wavefront is obtained by maximizing the feedback signal. Size of the sub-aperture is chosen to be small enough so that the diffraction-limited focus generated by this sub-aperture is substantially larger than the object itself (nominally by a factor of 2). We repeat the process for a sequence of sub-apertures. The optimized wavefront for each sub-aperture is then treated as a macro-mode. **b**, Next, we select a pair of adjacent sub-apertures and adjust the (relative) phase offset between their macro-modes to maximize reflection from the object. The optimized pair create a diffraction-limited focus that is approximately the size of the object itself. This process is repeated for all sub-aperture pairs and the optimal phase mask is obtained when this process finishes. **c**, By transmitting light through all the sub-apertures simultaneously (i.e., through the full aperture) with the optimized phase mask, a sharp, diffraction-limited focus is generated at the "center" of the object. This final focus is significantly smaller than the object and is raster scanned to obtain an image of the hidden object. The scales and angles shown in the illustration are different from our actual implementation.

and scannable optical focus at the object where the size of the focus is substantially smaller than the object itself. From optical geometry, we expect that the achievable focus spot size is related to the subtended illumination angular aperture at the wall, which depends on the magnification power of the imaging system, the distance between the wall and the hidden object, and the pixel size of the SLM/DMD. As the input pupil and the wall are image-forming conjugate planes, we can control the achievable spot size by controlling the aperture size on the wall. If the object is too large and it is unable to achieve the required de-magnification power in practice, gating methods can be utilized (Section 4.6). A fully opened pupil (corresponds to full aperture) in combination with the correct pattern displayed on the SLM should render a diffraction-limited focal spot on the object, as shown in Fig. 4.3c.

Mathematically, the relationship between the full aperture size at the wall and the focal spot size (characterized by its full width at half maximum, FWHM) is equal to $\frac{\lambda d}{s}$, where s is the lateral size of the full aperture, λ the wavelength of the light, and d the distance from the wall to the target. By leveraging the wall's optical memory effect [68] and superposing a suitable spatial phase ramp on the SLM phase pattern, we should be able to scan the focal spot across the object and its surrounding (Fig. 4.4). The determination of the correct SLM phase pattern is the central goal of UNCOVER. This is accomplished in UNCOVER through a 3-step process. In the following discussion, the (sub-) aperture refers to the illuminated area on the wall.

In Step 1, we partition the full aperture into Q sub-apertures. We select the size of the projected sub-apertures on the wall so that two adjacent, optimized sub-apertures would render a diffraction-limited focal spot at the object that is considerably larger than the object's size for the correct SLM phase pattern (the reason will become clear later). That is, one optimized sub-aperture would render a diffraction-limited spot that is substantially larger than the object itself (nominally by a factor of 2 laterally). In this case, it can be proven that the hidden object can be treated as a point (detailed requirement and its derivation can be found in the second part of Section 4.6).

In Step 2, we adjust the DMD to project one such sub-aperture. Our goal in this step is to find the correct SLM phase pattern which focuses light onto the object. We further note that the SLM and the DMD are part of the same image-forming conjugate set. To find the correct wavefront solution for one of the sub-apertures, phases of a relatively small number of SLM pixels are modulated such that the feedback signal measured by the detector (Fig. 4.3) is maximized. The determination of the right SLM phase pattern follows a fairly standard wavefront shaping strategy. In our experiments, we used a Hadamard pattern based search process [53, 55, 56, 59, 69], but we can expect most other feedback-based wavefront shaping strategies to work as well. We repeat this procedure for each of the Q sub-apertures. At the end of Step 2, we can treat the determined SLM phase solution for each sub-aperture as a macro-mode. Each macro-mode should focus light at the object with the focal spot size that is diffraction limited (approximately twice the size of the object itself).

In Step 3, we need to determine the correct phase relationship between the sub-aperture macro-modes to synthesize a full aperture, which generates a diffraction-limited focal spot at the object. To accomplish this, we start by opening up 2 adjacent sub-apertures and displaying their corresponding phase patterns on the SLM. One

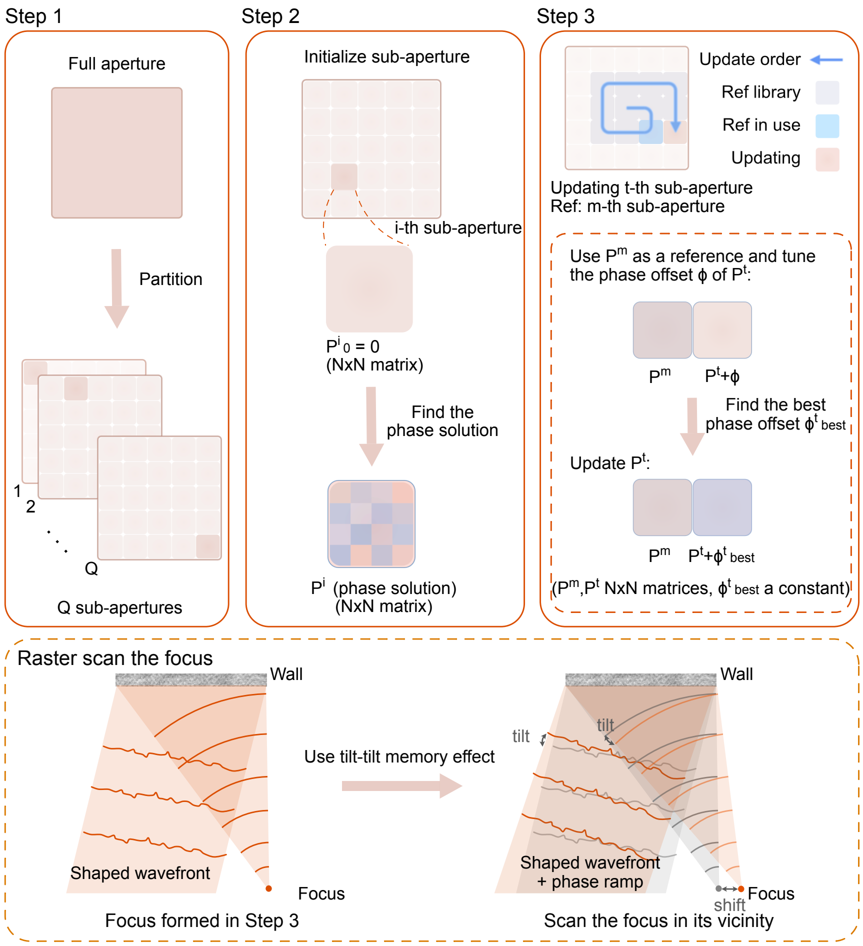

Figure 4.4: Step by step procedure for UNCOVER. Pupil segmentation is performed in step 1 such that the focus generated by two adjacent sub-apertures is smaller than the object. In step 2, phase solutions for each sub-aperture are optimized individually using iterative feedback-based wavefront shaping methods. In step 3, one of the sub-apertures (sub-aperture m in the figure) are used as the reference and a global phase offset is added to the other adjacent sub-aperture (sub-aperture t). By maximizing the feedback signal, the phase solution P^t for sub-aperture t is updated by adding the best phase offset Φ^t_{best}. Here, Q, N, i, m and t are integers. P^k denotes phase mask for the k-th sub-aperture, which is a N-by-N matrix. Once the focus is generated on the target, such focus can be raster scanned, with the help of tilt-tilt memory effect, in the surrounding of the object to generate an image.

of these patterns will serve as the reference and we will adjust the global phase offset of the other to maximize the return feedback signal from the object. We note here that the optimized pair should now render a diffraction-limited focal spot that is approximately the size of the object itself. More importantly, this focal spot is definitively no smaller than the target so the target can still be treated as a point.

Having found the correct global phase relationship for these two sub-apertures, we will then repeat the process with the next adjacent sub-apertures, and so forth. In our experiments, we perform this global phase determination process by selecting sub-apertures in a spiral sequence as shown in step 3 of Fig. 4.4. This process is akin to the standard feedback-base wavefront shaping procedure, except we are treating each sub-aperture as a macro-mode and simply adjusting each sub-aperture's global phase using pairs consisting of two adjacent sub-apertures, which confine the size of the focus in the optimization. We do not conduct this optimization for sub-apertures that are not adjacent to each other.

At the completion of Step 3, light from all sub-apertures should interfere constructively at the center of mass of the object's reflection function (derivations can be found in the second part of Section 4.6). By using the pairs in UNCOVER, we can find the relative phase offsets so that light from all sub-apertures will constructively interfere at the center of the target. Applying this optimal phase offset and using the full aperture, the rendered focus should be significantly smaller than the object. This ability to use the object as the guidestar and yet render a smaller-than-object and near-diffraction-limited focus at the object is a key innovation in UNCOVER. The focal spot is then scanned across the object by imposing a spatial phase ramp on the global phase pattern found in Step 2. The extent by which we can shift the focal spot is determined by the optical memory effect range associated with the wall and is related to the wall roughness [67, 68].

4.4 Experiment results of UNCOVER

Our proof-of-concept UNCOVER experiment setup is shown in Fig. 4.5. A continuous-wave laser diode (DJ532-40, Thorlabs) is used as the light source. The laser beam passes through two mirrors and a half-wave plate (HWP) and is coupled into a polarization maintaining fiber (PM fiber) for spatial filtering. The half-wave plate is used to align the polarization of the laser to the fast axis of the PM fiber and the mirrors are used to couple the light into the fiber. The filtered light exiting the fiber is expanded to fully cover the active area of SLM (Pluto NIR II, phase only SLM, Holoeye). The beamsplitter (BS) and the 4f system (depicted as L3 and L4 in Fig. 4.5) images the wavefront shaped light reflected by the SLM onto the DMD (DLP lightcrafter 6500 evaluation module, Texas Instrument). The zoom lens (L3; Canon EF-S 18-55 mm) is used to match the different pixel sizes on SLM and DMD. DMD is pixel-to-pixel conjugated to SLM and used to select the sub-aperture to use in the optimization step. A mirror and a second 4f system then project the DMD reflection onto the wall. And the reflection from the wall finally impinges on the object. In our experiment, we use a diffuse reflector (Thorlabs, DG10-1500-P01) as the wall proxy. A photomultiplier tube (PMT, H9306-03, Hamamatsu) and a condenser lens (Thorlabs, ACL50832U-A) are used to collect the light that returns from the object by way of diffuse reflection from the wall. The floor is covered with black papers to minimize the stray light and ensure the bright object is well-defined. Signals from both DMD and PMT are recorded using a data acquisition card (DAQ, PCI-MIO-16XE-10, National Instrument). After the UNCOVER wavefront optimization has been completed, a neutral density filter (ND) is inserted prior to raster scanning to prevent the light from saturating the PMT. The DMD's modulation determines the sub-apertures used in the optimization process. The number of sub-apertures (Q) and number of independent modes (N^2) in each sub-aperture will be specified for each experiment. The optimization process pipeline for UNCOVER can be found in the first part of Section 4.6.

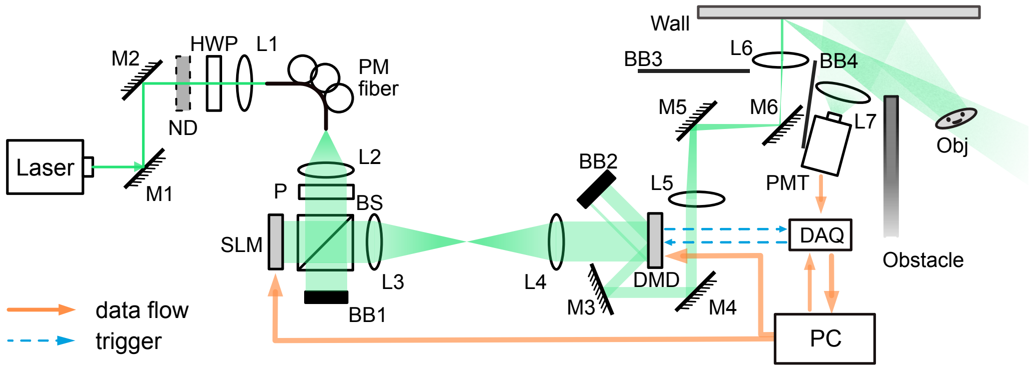

Figure 4.5: System setup of UNCOVER. Light is spatially filtered using the polarization maintaining fiber (PM fiber) and then gets expanded. The spatial light modulator (SLM) and digital micro-mirror device (DMD) are pixel-to-pixel conjugated using the 4f system consisted of L3 and L4. Then DMD is conjugated onto the wall by another 4f system (L5 and L6). DMD's spatial on-off modulation creates different sub-apertures. BB: beam block, BS: beam splitter, DAQ: data acquisition device, HWP: half-wave plate, L: lens, M: Mirrors, ND: neutral density filter, Obj: object, and PMT: photomultiplier tube. In our experiment, SLM/DMD is projected onto a reflective ground glass (wall proxy). Light passes through L6 is normal to the wall surface and the angle between this incident light and the scattered light reaching the target is around 32 degrees in our experiment. The object is also approximately normal to the incident light. BB3 and BB4 are used to minimize the stray light in our system. The floor is covered with black papers to ensure the bright object is well-defined and to minimize the stray light.

After we determine the correct wavefront solution for each sub-aperture (Step 2), we would then use sub-aperture pairs to find the phase offset among all the sub-apertures. The pre-defined sub-aperture pairs are loaded to DMD before the optimization to improve speed. We then update the relative phase between the pair in use. We repeat this process until all the relative phases are optimized. (Fig. 4.9) A tight focus is then generated when we actuate the DMD for full aperture transmission. We can then impose a linear spatial phase ramp on the SLM wavefront solution to raster scan the focus at the object. The scan range is determined by the tilt-tilt memory effect supported by the wall. In our experiments, our wall proxy has a tilt-tilt memory effect range of 2°, where the angle is measured when the intensity of the focus drops by half while using the largest synthetic aperture.

The first experiment is shown in Fig. 4.6. An object target was created by placing a patterned black mask (the nominal size of the target was 4 mm) on a diffuse reflector (Thorlabs, DG10-1500-P01). The object was placed 0.3 m away from the wall. The PMT was placed 0.35 m away from the wall. The object and the PMT's direct line of sight was blocked by an obstacle. The mean measurement time for each feedback signal is 40 ms in Step 2. In this experiment, $Q = 64$ sub-apertures and each with $N^2 = 64$ independent modes were used. After performing the UNCOVER optimization process, we were able to render a focus on the object (Fig. 4.6a3-c3). We then raster scanned the focus across the object and measured the PMT signal to render a UNCOVER image of the object. For each UNCOVER image pixel, we measured the PMT signal for 100 ms.

In this experiment, the measured focal spot size (FWHM) at the object is 0.55 mm. This measurement was acquired by replacing the object with a camera. The measured spot size is close to the diffraction limited spot size ($\sim$ 0.35 mm in this setting) — indicating that UNCOVER is indeed able to provide close to diffraction limited focusing. Additionally, the distance-to-resolution ratio was determined to be 550. For all three target shapes shown in Fig. 4.6, the measured focus spot size was relatively unchanged. It is additionally worth noting that the spot size is $\sim$ 7 times smaller than the targets — clearly demonstrating that UNCOVER is capable of generating a focus smaller than the object and is thus able to support imaging. The scanned UNCOVER images are shown in Fig. 4.6, demonstrating good correspondence to the target shapes.

The peak-to-background ratio (PBR) of the focus was also measured for all three shapes. The theoretical PBR should only scales as the number of SLM pixels used by UNCOVER. In this experiment, the theoretical PBR should be equal to 3220.

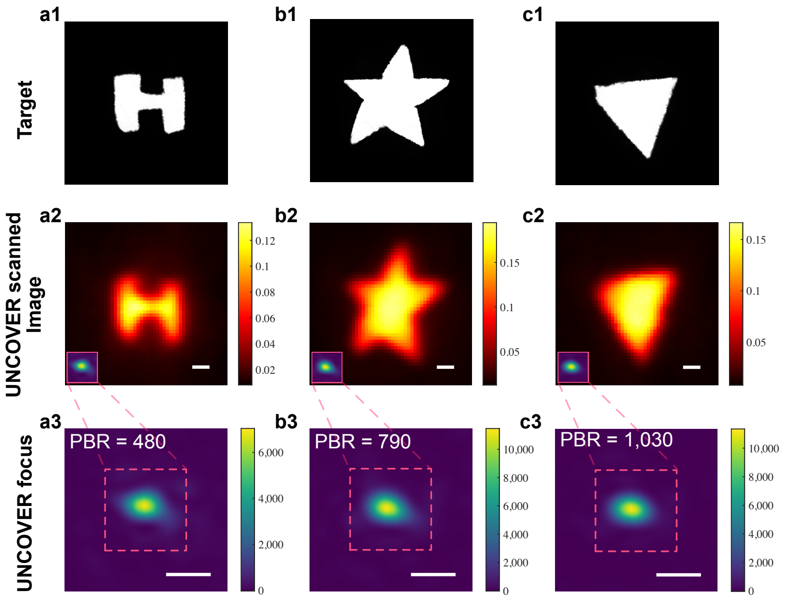

Figure 4.6: UNCOVER results using different targets. **a1**-**c1**, the ground truth of the target. **a2**-**c2**, results of UNCOVER. The inserted figure on the lower left of **a2**-**c2** is the focus in the dashed square of **a3**-**c3**, respectively. And it is adapted to have the same scale with respect to the scanned image. **a3**-**c3**, the image of the optimized foci. The measured peak-to-background ratio (PBR) is shown as well. The result is acquired using $Q = 64$ sub-apertures, with $N^2 = 64$ independent modes in each sub-aperture. A square of 9 by 9 pixels on SLM is treated as one mode. The total illuminated area is a square with side length of approximately 0.35 mm. Scale bars: 1 mm.

Interestingly, our measured PBR for the three shapes varies significantly, ranging from 480 to 1030. This variation and discrepancy from the theoretical prediction is attributable to the fact that our measurements are noisy. In future experiments, this issue can be mitigated by using a higher power light source and/or by reducing stray light. The degradation of our measured PBR due to noise is consistent with our observation that our lowest measured PBR came from the experiment with the H-shaped target. That target reflected the least light and, as such, gave us the lowest SNR to work with. To study the PBR one can achieve under different initial conditions, we performed additional simulations and the results can also be found in Section 4.6.

As UNCOVER generates a scannable optical focus at the object, it is possible for us to raster scan the generated focus beyond the object and image neighboring objects within the range supported by the wall's memory effect. This is potentially an advantage for UNCOVER for imaging dim objects in the presence of bright objects.

Bright objects and their associated higher noise can overwhelm the contributions of dimmer objects in standard NLOS method, leading to a diminished imaging dynamic range during image rendering. UNCOVER's ability to focus light on a dim object should allow us to exclude contributions from bright objects when we are attempting to image a dim object.

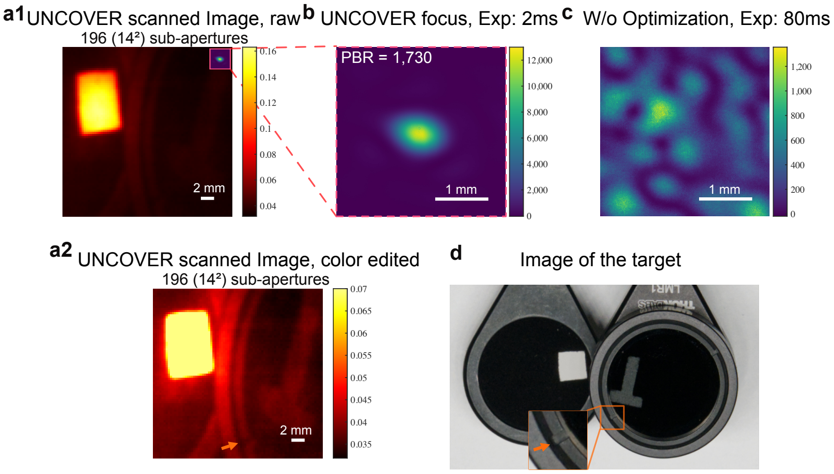

Figure 4.7: UNCOVER result of object with large reflectance variation. **a1**, the raw image from UNCOVER. Inset on the upper right of **a1** is adapted such that the image and the focus are on the same scale. **a2**, result with the maximum signal being capped at 0.07 to make the 'T' shape object more visible. **b**, the focus produced by UNCOVER. **c**, the speckle pattern before optimization. **d**, the target used in this experiment. The brighter square type object is 0.55 m away from the wall. An OD = 0.6 ND filter is placed in front of the 'T' shape object, which reduces its reflection coefficient to 6.3% when comparing with its original reflection coefficient. The arrows in **a2** and **d** show the resolved slot of the retaining ring. In this experiment, a square of side length ≈ 0.61 mm is illuminated when using the full aperture.

To study this, we prepared a composite object consisting of a square target placed next to the 'T' target. A ND filter was placed in front of the 'T' target so that the diffuse reflection from it is 16 times weaker than the square target's reflection. In this experiment, the object was placed 0.55 m from the wall. The PMT is 0.5 m away from the wall. Here, we used $Q = 196$ sub-apertures each with $N^2 = 64$ independent modes to achieve a higher lateral resolution and a more uniform background. The mean measurement time for generating one feedback signal in Step 2 is 600 ms. For each UNCOVER image pixel, we measured the PMT signal for 400 ms. From the result (Fig. 4.7 a1 and a2), the weaker target is clearly resolved in the scanned image. In fact, the return signal from the weaker object is 2.5 times higher than the null background. When both this ratio and the relative reflection difference between the weak and primary objects are taken into consideration, this implies the contrast range of our system is $\sim$ 40. In terms of resolution performance, the measured resolution here is 0.57 mm (FWHM) and the measured distance-to-resolution ratio is 970 in this experiment (the theoretical limit is $\sim$ 1430 for this experiment).

UNCOVER's ability to provide large contrast variation accommodation compares well with the performance of the state-of-the-art ToF based NLOS method, in which the intensity of the reconstruction artifact can be comparable to the primary object [46].

4.5 Discussion on UNCOVER

We demonstrated that UNCOVER can be used to generate a scannable diffraction limited focus that is significantly smaller than the target for NLOS imaging. UNCOVER deviates significantly from standard NLOS method in that it noninvasively forms a small scannable focus at the object.

A direct advantage of such an active focusing based NLOS imaging approach is that it is capable of performing diffraction limited imaging. In our experiment, we demonstrated a prototype system capable of generating a focus spot of size 0.57 mm at a distance of 0.55 m from the wall. This spot size compares favorably with the theoretical limit of 0.37 mm. Additionally, we experimentally demonstrated a distance-to-resolution ratio of 970 in our experiment. This performance specification significantly exceeds the best reported distance-to-resolution ratio of $\sim$ 100 in previous state-of-the-art NLOS experiments using no prior information of the wall [44–47].

Our experiments also demonstrated that it is possible to redirect the UNCOVER generated focus onto a dimmer object to image the dimmer object. In doing so, we can reduce the return signal from the brighter object and, in turn, allow us to better isolate the signal contribution from the dimmer object. Furthermore, we demonstrated a contrast range of 40 in our experiments, which compares well with standard NLOS methods which suffers from isolating signal contributions from dimmer object in post-processing. Our demonstration experiments were not optimized to maximize the image contrast dynamic range. In future experiments, it would be interesting to explore the bounds of this range, as well as to formalize a mathematical framework to understand its theoretical limit.

We note that our UNCOVER demonstration prototype was not optimized for speed. The process of UNCOVER optimization in our experiments was very much limited by the low laser power and the slow refresh rate of the SLM. Finding the correct wavefront solution to render a UNCOVER focus showed in Fig. 4.6 a2–c2 required a total time of $\sim$ 40 minutes. Generating the UNCOVER image in Fig. 4.6 a3–c3 took another 10 minutes. We expect this time can be significantly reduced by switching to high power lasers and faster wavefront shaping methods, such as those based on the use of DMDs, ferroelectric SLM, etc.

At first glance, it may appear that UNCOVER has violated one of the tenets of wavefront shaping — it is not possible to generate a focal spot that is smaller than the homogeneous guidestar/object [66] with linear feedback. UNCOVER overcomes this limitation by novelly exploiting the fact that the wall is a 2D scattering object and that it is possible to control the angular aperture on the wall itself. It is further worth noting that the singular and well-behaved focused spot is generated at the object reflection profile's center-of-mass, even if that location is not reflective at all (see Section 4.6).

UNCOVER's field-of-view (or lateral scan range) is limited by the tilt-tilt optical memory effect range of the wall. The wall proxy in our experiments afforded us a fairly large memory effect range (2 degrees). At a wall-to-object distance of 0.55 m, this corresponds to a nominal FOV of 4 cm laterally. We caution that this is a significant underestimation, as the focused peak can persist strongly (PBR $\gg$ 1) beyond several multiples of memory range using the model established by Osnabrugge et. al [67]. We additionally note that a rougher wall surface would provide a much narrower memory effect range (for example, 0.33 degrees for a piece of paper) [64, 68] . This topic requires a more complete study as UNCOVER's FOV is a complex function of the memory effect range, initial focus' PBR and hidden targets' reflection profile. We further note that the UNCOVER's FOV can be extended by optimizing for multiple foci at different locations, so that their total field-of-view will provide a much large constitutive coverage. To generate multiple foci, we may need to actively confine the spatial range of the photons that UNCOVER selectively detected. In this context, ToF gating of photons may be quite useful as a selection mechanism. Gating methods are also needed when there are multiple large, bright objects, as discussed in Section 4.6.

The angular scattering profile of the wall only impacts UNCOVER in its ability to collect photons effectively — a weakly scattering wall will reduce the overall transmission and UNCOVER would have to increase measurement time to compensate. The entire UNCOVER process of generating a focus spot is independent of the wall's scattering profile.

Finally, as with all other active illumination NLOS imaging methods, UNCOVER has a detection signal that drops off as a steep d^4 function (d^2 from illumination multipled by d^2 from light diffused back to the wall), where d is the distance from wall to target. This is a fundamental problem that confounds all active illumination NLOS imaging methods, and that ultimately provides a hard bound on the NLOS imaging range when light budget and imaging time are specified. UNCOVER has the potential to extend this range, in comparison to ToF NLOS methods. This is because ToF methods generally confine the area of detection in order to encode the shape of the target into the time differences of the arriving photons. This confinement restricts the photons that are used in the imaging process. In contrast, UNCOVER can fully utilize almost all the photons (except the one returns back to the full aperture) returning from the object via the wall. This boost in photon utilization rate can in turn be traded off for a longer NLOS imaging range. It is also interesting to note that this ability implies that UNCOVER detection signal is almost independent of the wall-to-detector distance — another signal advantage versus most ToF NLOS methods. We report a brief investigation on this point in Section 4.6. The UNCOVER signal dependency is worth a more extended and detailed future study.

In conclusion, we report a novel active focusing based NLOS imaging method that is able to generate a diffraction-limited smaller-than-guidestar focus for raster scanning the hidden object without any prior manipulation of the hidden scene. UNCOVER also broke new wavefront shaping ground by demonstrating that by controlled angular aperture on a scattering surface, it is possible to render a focused spot that is substantially smaller than the homogeneous guidestar itself. As with other active illumination NLOS imaging methods, there are still engineering challenges that need to be addressed for it to become a broadly useful NLOS imaging technique.

4.6 Supporting materials for UNCOVER

Optimization pipeline

To sync the digital micro-mirror device (DMD) and acquire signal coming from the photomultiplier tube (PMT), we chose to acquire the DMD output trigger and the signal from the PMT simultaneously. An example of the pipeline for our second step to find the optimal phase mask for each of the sub-apertures is shown in Fig. 4.8, where 9 sub-apertures were chosen for illustration.

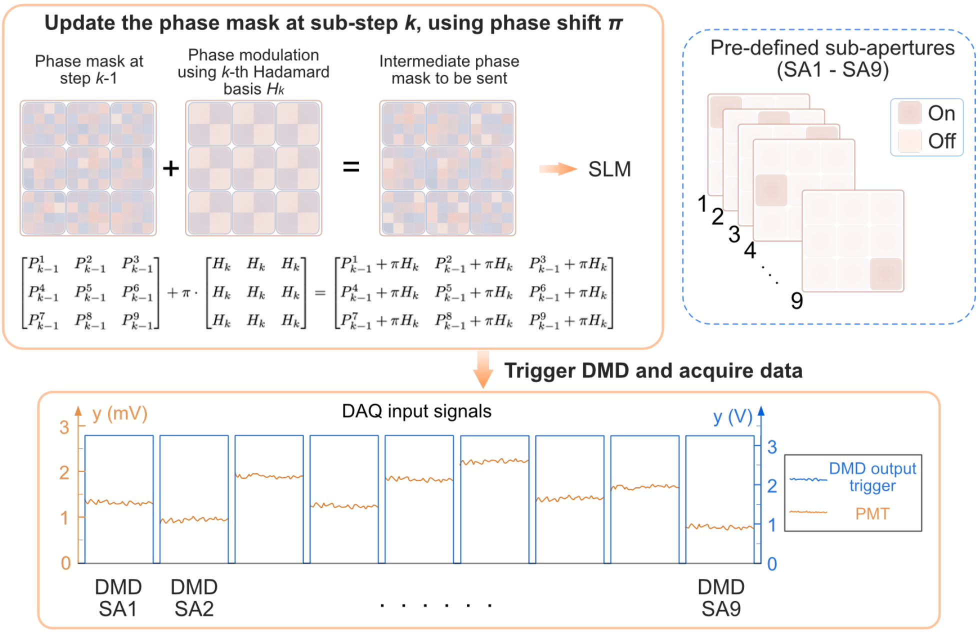

Figure 4.8: Pipeline for optimizing each of the sub-apertures in UNCOVER. In this figure, 9 sub-apertures are used for illustration. SLM is used in combination with a DMD for a higher optimization speed. DMD is applied to select which sub-aperture to use in the optimization. At sub-step k, the same phase modulation is added to all the optimized phase masks at step $k-1$. The newly generated phase mask is sent to SLM for updating. After SLM finishes updating, pre-defined sub-apertures are displayed sequentially by using the DMD. A small period of dark time is introduced to the DMD for discriminating signals from different sub-apertures. Signals within one of the blue window are averaged and the mean value is used in updating the phase mask. After finishing 4-step phase shift, the phase masks get updated and are used in step $k+1$. DAQ, data-acquisition device, P^i_k denotes the phase mask at sub-step k for sub-aperture i (P^i_k is a N-by-N matrix), and H_k denotes the k-th Hadamard basis (also a N-by-N matrix).

If we want to optimize all the sub-apertures, one way is to simply perform the iterative algorithm for all sub-apertures sequentially. In this case, if we have N^2 independent modes in each sub-aperture and in total Q sub-apertures, we need to update the SLM $4N^2 \times Q$ times using 4-phase shifting method (Hadamard basis based algorithm) [53, 55, 56, 59, 69]. This is inefficient as, in each sub-step k, only a small part of the SLM gets refreshed, leaving a great number of pixels unchanged. As optimizing one sub-aperture is independent of optimizing another, we use the same Hadamard pattern to modulate all the sub-apertures at each sub-step k. Thus, we can update all the phase masks by sending only one image to the SLM. We note that we still want to get the feedback signal with contribution from only one sub-aperture. Thus, we still need to block light from other sub-apertures when doing the measurement. This is done by using a much faster DMD (with a refresh rate of 9500 Hz for displaying binary patterns), which is simply used to select which sub-aperture to use during the optimizing process. In this case, patterns sent to SLM need only to be updated $4N^2$ times. This strategy is used in our experiment.

We first load the pre-defined sub-apertures (SA1 to SA9) into the buffer of DMD, the exposure time for all the patterns is defined at the beginning. A short period of dark time is also defined in order to discriminate signals when using different sub-apertures. All sub-apertures are initiated with zeros at the beginning. At sub-step k, the positive entries of the k-th Hadamard pattern are phase modulated using 0, $\pi/2$, π, and $3\pi/2$ [53, 55, 56, 59, 69].

Figure 4.8 shows the modulation step for phase modulation π. The intermediate phase mask is sent to SLM and signals from all the sub-apertures are acquired sequentially. After the signals for all the phase modulation are acquired, the phase mask is then updated. This step is repeated until the number of iterations reaches a predefined number.

If we write the feedback signal for SAi under the four different modulations as S_1^i, S_2^i, S_3^i, and S_4^i, respectively, the best phase at sub-step k is

$$\varphi_{\text{best}}^{i,k} = \arg\left[(S_1^i - S_3^i) + j\cdot(S_2^i - S_4^i)\right], \tag{4.1}$$

where $\arg(\cdot)$ takes the argument of a complex number, j is the unit imaginary number. Then, the phase mask of the i-th sub-aperture is updated in the following way:

$$P_k^i = P_{k-1}^i + \varphi_{\text{best}}^{i,k} H_k. \tag{4.2}$$

After the optimal phase solution for each of the sub-apertures has been obtained, we will start using pairs to optimize the relative phase offset among all the sub-apertures.

For tuning the phase offset among all the sub-apertures, we chose to update the phase offset along a spiral line, as shown in Step 3 of Fig. 4.4. Here, we showed a detailed example of the pairs when using 9 sub-apertures in total (Fig. 4.9). These pre-defined pairs are preloaded into the DMD for higher refresh rate. To update the k-th pair, the sub-aperture highlighted in blue, P^t, is chosen as the reference and keeps untouched in this step. Phase offset of 0, $\pi/2$, π, and $3\pi/2$ is added sequentially to the one to be updated (P^m), and generates an intermediate phase mask. This intermediate phase mask is uploaded to the SLM and feedback signals is

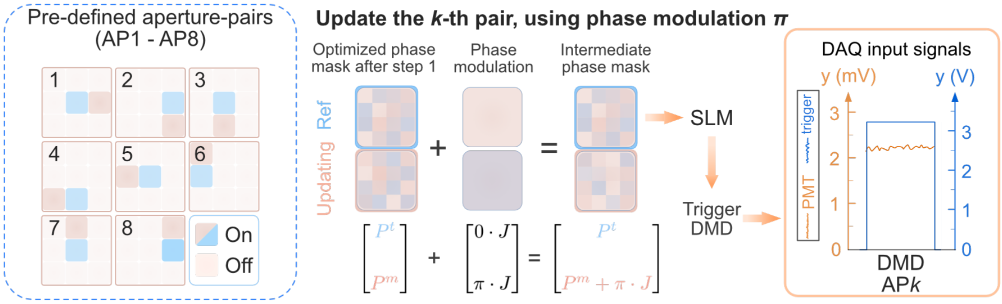

Figure 4.9: Using pairs to find the phase offset in UNCOVER. The blue color in the pre-defined sub-apertures denotes the reference in use, and the other highlighted sub-aperture is the one being updated. To update the k-th pair, phase offsets of 0, $\pi/2$, π, and $3\pi/2$ are added to the sub-aperture being updated. For each of the phase offset, the intermediate phase mask is sent to SLM. After SLM finishes refreshing, DMD is triggered to display the k-th aperture-pair, APk. The best phase offset is then calculated and added to the sub-aperture being updated using 4-step phase shifting method. P^t is the reference for the k-th pair (APk), P^m us the sub-aperture to be updated, and J is an matrix with all 1's whose dimension is the same as P^t and P^m.

then acquired by triggering the DMD to display the k-th aperture-pair (APk). Once the best phase offset that maximizes the feedback signal is obtained, this offset is directly added on P^m, which may then be used as a reference in the following steps.

If we write the feedback signals for updating the m-th sub-aperture (P^m) as $\hat{S}_1^m$, $\hat{S}_2^m$, $\hat{S}_3^m$, and $\hat{S}_4^m$, respectively, the optimal phase offset for P^m is:

$$\Phi_{\text{best}}^m = \arg\left[(\hat{S}_1^m - \hat{S}_3^m) + j \cdot (\hat{S}_2^m - \hat{S}_4^m)\right]. \tag{4.3}$$

And the phase mask of the m-th sub-aperture is updated:

$$P_{\text{updated}}^m = P^m + \Phi_{\text{best}}^m. \tag{4.4}$$

If the phase offsets have been obtained for all the sub-apertures, a sharp focus is generated when we open up all the sub-apertures and apply the optimized phase mask. This focus is then raster-scanned to get an image of the target.

Derivation of UNCOVER

In the following section, we provide a framework for analyzing UNCOVER. We will start from the simplest example where we want to align two phasors to maximize the feedback signal. Then, we will extend the approach used in this toy example to find the solution of the maximizing problem for one of the sub-apertures. The intuition comes from the idea that when the size of the sub-aperture is confined in a way such that the speckle size (which is approximately equal to the focus size after optimization) is larger than the object, the object can be treated as a fixed point (i.e., the location of the point is independent of the selection of sub-aperture).

To synthesize all the sub-apertures, we want the foci generated by the sub-apertures to interfere constructively at the same point (thus enable us to generate a sharp focus). We will see that, with constraint on the size of the sub-aperture pairs applied in the optimization step, the object acts as a fixed point throughout the optimization. We will show that a final focus is formed on the center of mass of the object function, which will be defined later on.

The maximizing problem for one of the sub-apertures

Before we dive into the details, we will first state the assumption of UNCOVER and derive the constraints on the object such that 3D objects can be approximated as 2D objects. The derivations from Proposition 4.1 to Eq. 4.21 prove that 3D objects can be treated as 2D objects with constraint on the axial extent of the 3D objects. After we have proven that 2D approximation is feasible, we will conduct our analysis and derive the result based on 2D objects.

Assumption 4.1 *Assume the distance between the wall and the object is large, so that light emitting from each mode of one sub-aperture can be approximated as a plane wave over the (volume of) the hidden objects.*

Using the above assumption, the speckle incidents on the object caused by light interfering from one of sub-apertures can be written as

$$\sum_n A_n(\mathbf{k_n}) e^{i(\mathbf{k_n}\mathbf{r}+\phi_n)}, \tag{4.5}$$

where $\mathbf{k_n}$ is the wave vector for input mode n, A_n the amplitude of input mode n, $\mathbf{r}$ the spatial location and ϕ_n the phase of that mode.

Let us first consider a 3D object function $O(\mathbf{r}) = O(x, y, z)$, the signal received by the detector, E, is given by

$$E = \int_O I(\mathbf{r}) \cdot O(\mathbf{r}) d\mathbf{r} = \iiint I(x, y, z) \cdot O(x, y, z) dx dy dz. \tag{4.6}$$

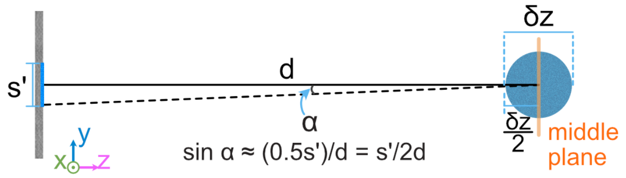

Figure 4.10: Variables in proposition 4.1. The angles shown here are for illustration and do not correspond to the angles in our experiments.

Proposition 4.1 *Under assumption 4.1, let us further assume that the size of the sub-aperture, denoted by s', is much less than the distance d between the wall and the object. The illumination pattern is approximately independent of z, given the axial extent δz satisfies the following constraint:*

$$\delta z \leq \min\left\{0.2d, \frac{1}{10} DOF_{sub\text{-}aperture}\right\}, \tag{4.7}$$

where $DOF_{sub\text{-}aperture}$ is the depth of field of the sub-aperture, and is defined using the formula of DOF in microscopy: $DOF_{sub\text{-}aperture} = \frac{\lambda}{NA^2}$.

Proof: Based on the assumption in Proposition 4.1, we have (α is the one-half angular aperture):

$$\mathrm{NA}' := \mathrm{NA}_{\text{sub-aperture}} = n_{\text{air}} \cdot \sin\alpha = \sin\alpha \approx \frac{s'}{2d} \ll 1, \tag{4.8}$$

and

$$\delta z \leq \frac{1}{10}\mathrm{DOF}_{\text{sub-aperture}} = \frac{1}{10}\frac{\lambda}{\mathrm{NA}'^2} = \frac{2\lambda d^2}{5s'^2}. \tag{4.9}$$

Note that Proposition 4.1 is based on Assumption 4.1, so we want the direction of $\mathbf{k_n}$ to remain the same over δz.

Let us first assume $\frac{\delta z}{2} \ll d$. To investigate what the constraint is, we will write the k-vector of mode n at axial position z' as $\mathbf{k_n}(z')$ for now (z' is some arbitrary value here). The direction change of $\mathbf{k_n}(z')$ can be characterized by the change in the polar angle $\alpha_n(z')$ of $\mathbf{k_n}(z')$, as the azimuthal angle in spherical coordinates is independent of z in Cartesian coordinates. If the tolerance of the absolute rate of change in $\alpha_n(z')$ is set to 0.1 when $z' \in [d - \frac{\delta z}{2}, d + \frac{\delta z}{2}]$, we require (we have

$\delta z \ll 2d$):

$$\left|\alpha_n(d) - \alpha_n(d \pm \frac{\delta z}{2})\right| \approx \left|\frac{s}{2d} - \frac{s}{2d \pm \delta z}\right| = \left|\frac{s}{2d} - \frac{s}{2d}\frac{1}{(1 \pm \frac{\delta z}{2d})}\right|$$
$$\approx \frac{s}{2d}\left|1 - (1 \mp \frac{\delta z}{2d})\right| \leq 0.1\alpha_n(d) \approx \frac{s}{20d} \tag{4.10}$$
$$\Rightarrow \quad \delta z \leq 0.2d.$$

To summarize, we get

$$\delta z \leq \min\left\{0.2d, \frac{2\lambda d^2}{5s'^2}\right\}. \tag{4.11}$$

If the above is satisfied, the Assumption 4.1 holds, and $\mathbf{k_n}$ is independent of z. Thus, in our following derivation, $\mathbf{k_n}$ can be regarded as a constant vector. The E-field E_f of the illumination pattern at $(x, y, d + \Delta z)$ is:

$$E_f(x, y, d + \Delta z) = E_f(\mathbf{r} + \Delta\mathbf{z}) = \sum_n A_n(\mathbf{k_n})e^{i(\mathbf{k_n}\cdot(\mathbf{r}+\Delta\mathbf{z})+\phi_n)}$$
$$= \sum_n A_n(\mathbf{k_n})e^{i(\mathbf{k_n}\mathbf{r}+\phi_n)}e^{ik_{n,z}\Delta z}. \tag{4.12}$$

Let us focus on the factor $e^{ik_{n,z}\Delta z}$. For the z-component of $\mathbf{k_n}$, we have (α_n is the polar angle of $\mathbf{k_n}$ in spherical coordinates):

$$k_{n,z} = |\mathbf{k_n}|\cos(\alpha_n) = \frac{2\pi}{\lambda}\sqrt{1 - \sin^2(\alpha_n)} \approx \frac{2\pi}{\lambda} - \frac{\pi}{\lambda}\sin^2(\alpha_n), \tag{4.13}$$

where $\alpha_n \in [-\alpha, \alpha]$, as α is the one-half angular aperture. Thus, we have

$$|k_{n,z} - k_{m,z}| \leq \max_n k_{n,z} - \min_m k_{m,z}$$
$$= \frac{2\pi}{\lambda} - \left(\frac{2\pi}{\lambda} - \frac{\pi}{\lambda}\sin^2(\alpha)\right) = \frac{\pi}{\lambda}\sin^2(\alpha). \tag{4.14}$$

The absolute phase retardation between input mode n and input mode m over Δz is bounded above by:

$$|k_{n,z}\Delta z - k_{m,z}\Delta z| = |k_{n,z} - k_{m,z}|\Delta z \leq \frac{\pi}{\lambda}\sin^2(\alpha)\Delta z. \tag{4.15}$$

Plug in Eq. 4.11 and use the fact that $\Delta z \in [-\frac{\delta z}{2}, \frac{\delta z}{2}]$,

$$\begin{aligned} |k_{n,z} - k_{m,z}|\Delta z &\leq \frac{\pi}{\lambda}\sin^2(\alpha)\Delta z \\ &\leq \frac{\pi}{\lambda}\sin^2(\alpha)\frac{\delta z}{2} \\ &\leq \frac{\pi}{\lambda}\sin^2(\alpha)\cdot\frac{1}{20}\frac{\lambda}{\mathrm{NA}'^2} = \frac{\pi}{20}. \end{aligned} \tag{4.16}$$

That is, the maximum path length difference over the axial extent of the 3D object function is $\frac{\lambda}{40}$. (Note that this is a strict criterion and can be eased.) Such a small phase retardation between different input modes can be ignored:

$$e^{ik_{n,z}\Delta z} \approx e^{ik_{m,z}\Delta z} = e^{i\theta_0}, \qquad \forall n, m. \tag{4.17}$$

Then, we have

$$\begin{aligned} E_f(x, y, d + \Delta z) &= \sum_n A_n(\mathbf{k_n})e^{i(\mathbf{k_n}\mathbf{r}+\phi_n)}e^{ik_{n,z}\Delta z} \\ &\approx e^{i\theta_0}\sum_n A_n(\mathbf{k_n})e^{i(\mathbf{k_n}\mathbf{r}+\phi_n)} = e^{i\theta_0}E_f(x, y, d). \end{aligned} \tag{4.18}$$

The intensity $I(x, y, z + \Delta z)$ is given by:

$$\begin{aligned} I(x, y, d + \Delta z) &= E_f(x, y, d + \Delta z)\cdot E_f^*(x, y, d + \Delta z) \\ &= E_f(x, y, d)\cdot E_f^*(x, y, d) = I(x, y, d). \end{aligned} \tag{4.19}$$

So, within the range $z \in [d - \frac{\delta z}{2}, d + \frac{\delta z}{2}]$, the illumination pattern is independent of z and can be written as a function of (x, y) only:

$$I(\mathbf{r}) = I(x, y, z) = I(x, y). \tag{4.20}$$

The signal E can be simplified:

$$\begin{aligned} E &= \iiint I(x, y, z)\cdot O(x, y, z)dxdydz \\ &= \iint dxdy\int_{d-\frac{\delta z}{2}}^{d-\frac{\delta z}{2}} I(x, y)\cdot O(x, y, z)dz \\ &= \iint dxdyI(x, y)\cdot\int_{d-\frac{\delta z}{2}}^{d-\frac{\delta z}{2}} O(x, y, z)dz \\ &= \iint I(x, y)\cdot O'(x, y)dxdy, \end{aligned} \tag{4.21}$$

where $O'(x, y) := \int O(x, y, z)dz$ is the projection of $O(x, y, z)$ onto the xy-plane.

Thus, we prove Proposition 4.1, which means a 3D object is equivalent to a 2D object when the assumptions in the proposition hold. Note that more sophisticated maps can be used for the 3D to 2D transformation, such as using cones to map the 3D objective function to a particular xy-plane. More works can be done in this direction.

In the following derivation, we will assume a 2D object function is used, otherwise it can be projected onto a 2D plane by using Proposition 4.1. And we are interested only in the 2D coordinates that involve (x, y).

In iterative wavefront shaping, we only modulate the phase of each input mode, and the output field with the phase modulation applied can be written as:

$$\sum_n A_n(\mathbf{k_n}) e^{i(\mathbf{k_n r} + \phi_n + p_n)}, \tag{4.22}$$

where p_n is the phase modulation we applied, $\mathbf{k_n} = [k_x, k_y]^T$ and $\mathbf{r} = [x, y]^T$. We define the maximum length $l_{\max}$ of an object A as:

$$l_{\max} = \max\{|\mathbf{r_1} - \mathbf{r_2}| : A(\mathbf{r_1}) \neq 0, A(\mathbf{r_2}) \neq 0\}, \tag{4.23}$$

where $|\cdot|$ is Euclidean norm.

Assumption 4.2 *Assume the maximum length of the object O, l_{max}, satisfies $\forall n, m : |\mathbf{k_n} - \mathbf{k_m}| l_{max} < \pi$ (the smallest period of the fringes is larger than l_{max}).*

The total energy detected $E(\mathbf{p})$ is given by:

$$\begin{aligned}
E(\mathbf{p}) &= \int_O O(\mathbf{r}) I(\mathbf{r}) d\mathbf{r} \\
&= \int_O O(\mathbf{r}) \left(\sum_n A_n(\mathbf{k_n}) e^{i(\mathbf{k_n r} + \phi_n + p_n)} \right) \cdot \left(\sum_m A_m(\mathbf{k_m}) e^{-i(\mathbf{k_m r} + \phi_m + p_m)} \right) d\mathbf{r} \\
&= \int_O O(\mathbf{r}) \left[\sum_n A_n^2(\mathbf{k_n}) + 2 \sum_{n<m} \Re \left(A_n(\mathbf{k_n}) A_m(\mathbf{k_m}) e^{i(\Delta \mathbf{k_{nm} r} + \Delta\phi_{nm} + \Delta p_{nm})} \right) \right] d\mathbf{r} \\
&\propto c + \int_O O(\mathbf{r}) \sum_{n<m} \Re \left(A_n(\mathbf{k_n}) A_m(\mathbf{k_m}) e^{i(\Delta \mathbf{k_{nm} r} + \Delta\phi_{nm} + \Delta p_{nm})} \right) d\mathbf{r},
\end{aligned} \tag{4.24}$$

where $\mathbf{p} = [p_1, p_2, \cdots, p_n]^T$, $\Delta\mathbf{k_{nm}} = \mathbf{k_n} - \mathbf{k_m}$, $\Delta\phi_{nm} = \phi_n - \phi_m$, $\Delta p_{nm} = p_n - p_m$, $\Re(\cdot)$ takes the real part of a complex number, $c \in \mathbb{R}$ some constant independent of

p, and $O(\mathbf{r})$ is the fraction of light being diffused by the object at **r** (in other words, the object function).

Question What is the best **p** that maximizes the integral?

In the following section, we will see that maximizing the total energy detected is equivalent to maximizing the intensity over a certain point, which is the same as using a point guidestar.

Solution

To find out the solution for Eq. 4.24, we start by considering the interference of two phasors $A_n(\mathbf{k_n})e^{i(\mathbf{k_n r}+\phi_n+p_n)}$ and $A_m(\mathbf{k_m})e^{i(\mathbf{k_m r}+\phi_n+p_m)}$. The interference of the two point sources results in a sinusoidal pattern on the object. By studying this simple case, we will know at what point the integral is maximized. If, by some happenstance, the optimal point is the same for any two phasors we choose, the global maximum will also be achieved at that point. And we will shortly see that this is indeed the case. We will first define the center of mass of the object's reflection function $O(\mathbf{r})$: (we will see, after all these derivations, that this is the point that all phasors are aligned.)

Definition 4.1 *Define the center of mass C_m of an object O as:*

$$C_m = \frac{\int_O O(\mathbf{r})\mathbf{r}d\mathbf{r}}{\int_O O(\mathbf{r})d\mathbf{r}}. \tag{4.25}$$

Without loss of generality, let us focus on $\Delta\mathbf{k_{nm}}$, and set up a Cartesian coordinate such that the direction of the positive x-axis is the same as the direction of $\Delta\mathbf{k_{nm}}$ ($k := k_x = |\Delta\mathbf{k_{nm}}|$), and let the origin be C_m. Given $\mathbf{k_n}$ and $\mathbf{k_m}$, we have:

$$\begin{aligned}
&\int_O O(\mathbf{r})\Re\left(A_n(\mathbf{k_n})A_m(\mathbf{k_m})e^{i(\Delta\mathbf{k_{nm}}\mathbf{r}+\Delta\phi_{nm}+\Delta p_{nm})}\right)d\mathbf{r}\\
&\qquad = A_n(\mathbf{k_n})A_m(\mathbf{k_m})\int dx\int O(x,y)\Re\left(e^{i(kx+\Delta\phi_{nm}+\Delta p_{nm})}\right)dy\\
&\qquad = a_{nm}\int \Re\left(e^{i(kx+\Delta\phi_{nm}+\Delta p_{nm})}\right)\left(\int O(x,y)dy\right)dx\\
&\qquad = a_{nm}\int f(x)\cos(kx+\Delta\phi_{nm}+\Delta p_{nm})dx\\
&\qquad = a_{nm}\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\Delta\phi_{nm}+\Delta p_{nm}),
\end{aligned} \tag{4.26}$$

where $f(x) = \int O(x,y)dy \geq 0$, $a_{nm} = A_n(\mathbf{k_n})A_m(\mathbf{k_m}) > 0$ (For the last step of Eq. 4.26, note that we have $\forall n, m : |\mathbf{k_n} - \mathbf{k_m}|l < \pi$ based on Assumption 4.2). Also, as we define the origin to be C_m, we have:

$$\int xf(x)dx = 0. \tag{4.27}$$

Let us study this function:

$$F(k,\alpha) = \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha)dx. \tag{4.28}$$

Lemma 4.1 $\forall f(x), \exists \alpha : \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha)dx \geq 0$

Proof: Assume $\forall f(x), \alpha : \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha)dx < 0$. Let $\alpha' = \alpha + \pi$ we have:

$$\begin{aligned}\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha')dx &= \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha+\pi)dx \\ &= -\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha)dx > 0. \end{aligned}\tag{4.29}$$

This leads to a contradiction. Thus, we have that for all $f(x)$, there exists α such that

$$\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha)dx \geq 0. \tag{4.30}$$

As we want to maximize $F(k,\alpha)$, let us confine the range of α such that $\forall \alpha \in S : F(k,\alpha) \geq 0$, where S is the constraint set:

$$S \coloneqq \{\forall \alpha \in S : F(k,\alpha) = \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha)dx \geq 0\}. \tag{4.31}$$

Assume F and its first and second order of partial derivatives are all continuous (it satisfies Leibniz's rule), let us find the partial derivative of $F(k,\alpha)$ with respect to α:

$$\begin{aligned}\frac{\partial F(k,\alpha)}{\partial \alpha} &= \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\frac{\partial \cos(kx+\alpha)}{\partial \alpha}dx \\ &= -\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\sin(kx+\alpha)dx; \qquad \alpha \in S \end{aligned}\tag{4.32}$$

$$\begin{aligned}\frac{\partial^2 F(k,\alpha)}{\partial\alpha^2} &= -\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\frac{\partial \sin(kx+\alpha)}{\partial\alpha}dx \\ &= -\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha)dx \le 0; \qquad \alpha \in S. \qquad (4.33)\end{aligned}$$

The second order partial derivative is non-positive. Thus, to maximize $F(k,\alpha)$, we simply need to solve for this equation:

$$\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\sin(kx+\alpha)dx = 0. \qquad (4.34)$$

Note that if we want to have a well defined focus (all the phasors are optimized such that they are all in phase at a certain point), we want Eq. 4.34 holds for all k at a certain $x_0 = \frac{\alpha}{k}$.

Let us look at a extreme case: $k \longrightarrow 0$. In this case, we have

$$\lim_{k\to 0}\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\sin(kx+\alpha)dx = \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\sin(\alpha)dx = 0. \qquad (4.35)$$

We know that $f(x) = \int O(x,y)dy \ge 0$ (light energy returned to detector cannot be negative), so to solve Eq. 4.35 we get

$$\sin(\alpha) = 0 \Rightarrow \alpha = 0. \qquad (4.36)$$

In the above equation, we limit $\alpha \in (-\pi,\pi] \cap S$. Thus, another possible solution $\alpha = \pi$ is dropped based on the fact that

$$\lim_{k\to 0}\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(kx+\alpha)dx = \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\cos(\pi)dx = -\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)dx < 0. \quad (4.37)$$

Hence it does not belong to the constraint set $\alpha = \pi \notin S$.

As Eqs. 4.34 and 4.36 need to hold for all k, we need the following equation holds

$$\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\sin(kx)dx = 0. \qquad (4.38)$$

Condition 4.1 *The object $O(x,y)$ has point symmetry with respect to its center of mass, which means $O(x,y) = O(-x,-y)$. And l_{max} is less or equal to $\frac{1}{\max|\Delta\mathbf{k}|}$.*

Using the above condition, we have

$$f(-x) = \int O(-x, y)dy = \int O(x, y)dy = f(x). \tag{4.39}$$

We can easily verify that Eq. 4.38 holds in this case.

For a more general case in which the object is of arbitrary shape, let us use Taylor expansion to rewrite the left hand side of Eq. 4.38. This gives

$$\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\sin(kx)dx = \int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\left(kx - \frac{(kx)^3}{3!} + \frac{(kx)^5}{5!} - \cdots\right)dx. \tag{4.40}$$

Condition 4.2 *The maximum length* l_{max} *of the object* O *is much smaller than* $\frac{1}{\max|\Delta\mathbf{k}|}$

Using condition 4.2, Eq. 4.40 becomes

$$\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} f(x)\sin(kx)dx \approx k\int_{-\frac{\pi}{k}}^{\frac{\pi}{k}} xf(x)dx = 0. \tag{4.41}$$

If we want the error of the approximation $\sin(kx) \approx kx$ smaller than 10^{-3}, we need $\frac{(\max|\Delta\mathbf{k}|l_{\max})^3}{3!} < 10^{-3}$. That is, we want $l_{\max} < \frac{0.18}{\max|\Delta\mathbf{k}|}$.

Conclusion and connection to UNCOVER

If Eq. 4.38 holds, let us go back to the last term of Eq. 4.24:

$$\begin{aligned}\int_O O(\mathbf{r})\sum_{n<m}\Re\left(A_n(\mathbf{k_n})A_m(\mathbf{k_m})e^{i(\Delta\mathbf{k_{nm}}\mathbf{r}+\Delta\phi_{nm}+\Delta p_{nm})}\right)d\mathbf{r}\\ \leq \int_O O(\mathbf{r})\sum_{n<m}\Re\left(A_n(\mathbf{k_n})A_m(\mathbf{k_m})e^{i(\Delta\mathbf{k_{nm}}\mathbf{r})}\right)d\mathbf{r}.\end{aligned} \tag{4.42}$$

Thus, we obtain

$$\underset{\mathbf{p}}{\operatorname{argmax}}\, E(\mathbf{p}) = -\phi + p_0, \tag{4.43}$$

where p_0 is a constant and $\phi = [\phi_1, \phi_2, \cdots, \phi_n]^T$.

The sufficient condition that guarantees we will have a well defined focus is given by Condition 4.1 or Condition 4.2 under Assumption 4.1 and Assumption 4.2. This indicates that, under certain conditions, if the speckle size generated by one sub-aperture is larger than the object, the object can be treated as a point.

These two conditions basically mean that if the object is centrosymmetric or the maximum length of the object is smaller than the smallest period produced by the sub-aperture, then the algorithm described in the beginning would give us a focus. And the center of the focus coincides with the center of mass (C_m, Definition 4.1) of the object function. The intuition for this is that when the period is larger than the object (it will not be able to resolve the object), the hidden object can be treated as a point.

Coming back to UNCOVER, one wants a tight focus that can be raster scanned over the object itself. To realize a tight focus, we want all the phasors for the full aperture are aligned at a certain point.

Recall that we have derived that if size of the aperture is carefully selected, the hidden object can be treated as a point source at C_m. That is, if the size the illuminated area (aperture) at any instant in the optimization algorithm satisfy a certain constraint (e.g., the focus size of the optimized aperture is larger than the object), a focus can be formed at C_m, the position of the equivalent point guidestar.

To optimize the phase masks in Step 2, the sub-apertures are designed such that they are small enough to satisfy the aforementioned constraints (Note that this is not the ultimate requirement on the size of sub-apertures, we will discuss a further constraint in the following two paragraphs). With the constraints being satisfied, only a single focus with its center at C_m is obtained for each sub-aperture in Step 2.

The next question is how to synthesize those sub-apertures so that they produce one tight focus on the object. If a tight focus is desired, we want the hidden object to act like a point guidestar when adjusting the phase offset among the sub-apertures. That is, we want the constraints on the size of the aperture to be satisfied at any instance throughout Step 3 of the optimization algorithm.

The constraint on the size of aperture that is simultaneously used in Step 3 is satisfied by using pairs. We designed the sub-apertures such that when opening any two adjacent sub-apertures, the speckle size is still larger than the object size (which meets the constraints stated above). Following the result we have in the previous subsection, the hidden object serves as a point. Let us focus on what happens to this point. If we write $\tilde{A}_i$ and $\tilde{\gamma}_i$ as the amplitude and the phase of the focus of i-th optimized sub-aperture, respectively. Maximizing the signal scattered by the hidden object over the phase offset $\tilde{\omega}$ is equivalent to

$$\underset{\tilde{\omega}}{\operatorname{argmax}}\, |\tilde{A}_t e^{i\tilde{\gamma}_t} + \tilde{A}_m e^{i\tilde{\gamma}_m + i\tilde{\omega}}|^2, \tag{4.44}$$

where t denotes the reference sub-aperture and m denotes the sub-aperture to be updated in the k-th pair. Note that this is true only when the hidden target is equivalent to a point guidestar.

The optimal solution is given by

$$\tilde{\gamma}_m + \tilde{\omega}_{\text{best}} = \tilde{\gamma}_t. \tag{4.45}$$

Thus, using Eq. 4.4, the phase of the updated m-th sub-aperture is

$$\tilde{\gamma}_m^{\text{updated}} = \tilde{\gamma}_m + \tilde{\omega}_{\text{best}} = \tilde{\gamma}_t. \tag{4.46}$$

That is, all the sub-aperture in the reference library share the same phase at C_m. Thus, when the optimization is performed over all sub-aperture pairs, light from all sub-aperture interferes constructively at C_m. Thus, after Step 3 is finished, a tight focus is generated at C_m when the full-aperture is used (Note that the phase masks are applied directly and no optimization is needed).

Discussion on the algorithm

To use UNCOVER, we want the sub-aperture pairs to satisfy a certain constraint, which is related to the size of the hidden objects. Thus, an estimate of the hidden objects' size is needed. However, it is worthwhile to point out that, such prior knowledge is unnecessary if UNCOVER is performed multiple times with different sub-aperture sizes in each trial. In Fig. 4.14, it is shown that UNCOVER scanned images converge when UNCOVER generates only one tight focus (which means that the size satisfies the requirement in UNCOVER). So, we can use the similarity of the UNCOVER scanned images to identify if the constraint is satisfied or not. Other metric can be use as well: One can use ToF method to measure the distance that the returned photons travelled. If UNCOVER generates a single focus, there is only one strong peak showing up in the measurement. We expect many other detection methods may serve as the metric as well.

Proposition 4.1 shows 3D objects can be treated as 2D objects via projection, if the assumption in the proposition is satisfied. A natural question to ask is how strong these assumptions are. In other words, how small the axial extent should be in order to enable such projection. Here, we will give some numbers based on our second experiment (Fig. 4.7). In the second experiment, the side length of the full aperture is 0.6 mm and the distance from the wall to the object is $d = 0.55$ m. Each sub-aperture pair has a maximum side length $s' = \dfrac{0.6\,\text{mm}}{14} \times 2 = 86\,\mu\text{m}$. Thus, the upper bound for the axial extent δz of the object is (Eq. 4.11)

$$\delta z \leq \min\left\{0.2d, \frac{2\lambda d^2}{5s'^2}\right\} = \min\left\{0.2 \times 0.55, \frac{2 \times 532 \times 10^{-9} \times 0.55^2}{5 \times (86 \times 10^{-6})^2}\right\}$$
$$= \min\{0.11, 8.7\} = 0.11\,\text{m}. \tag{4.47}$$

Thus, the maximum axial extent is mainly dependent on the distance between the wall and the object.

If objects in the scene are distributed at significantly different distances, filtering methods need to be applied to UNCOVER. For instance, time-resolved measurement can be combined with UNCOVER to confine the axial range of the filtered signal. With this filtering, the effective axial extent of the object function can be confined in order to satisfy the condition in Proposition 4.1. Without filtering, UNCOVER may have degraded performance in generating a tight focus or even fail, depending on the specific settings. It is worthwhile to point out that the tight focus formed with filtering can be scanned over larger distances by superposing a defocus phase term over the phase solution. This range is not limited by the constraint on the axial extent (Eq. 4.11).

In UNCOVER, it is required that the focus size in optimizing phase offset among the pairs to be smaller than the hidden object. As the numerical aperture of the focus in such optimization depends on the size of the sub-apertures and the distance between the wall and the hidden target (Eq. 4.8), proper demagnifying system might be required. If there are multiple large objects (or a single large object), it might be tricky/impossible to de-magnify the sub-apertures such that they meet the constraints in UNCOVER. In this case, other techniques should be used to filter the returned signal. Follow the same logic in confining the axial extent of the effective object function, time gated/coherence gated measurement can be used to restrict the effective size of the objects such that the size is small and fulfills the constraints in UNCOVER. We also note that the scan range in UNCOVER is related to the memory effect range associated with the wall [68]. For object that is larger than this scan range, the field of view (FOV) needs to be increased in UNCOVER in order to image the entire object. This can be done by optimizing multiple foci at different positions. By stitching the FOVs associated with each individual focus, the overall FOV can be expanded, as stated in the Discussion part of our main manuscript.

We notice there is another paper uses segmented pupil to perform adaptive optics (AO) that looks similar to UNCOVER [70]. This AO method assumes that light from each segment of the pupil can form a focus (with unknown deflection and phase offset). This focus is raster scanned and the scanned image is compared with the reference image to now the deflection. In UNCOVER, however, we do not have such focus to begin with, an reference image is infeasible as well. The focus generated by one optimized sub-aperture is larger than the object. Thus, image of the hidden object cannot be formed by raster-scan of this focus. In optimizing the phase offset, the AO method use a single reference and no constraint is needed with respect to the distance between the two pupil segments in the optimization. In contrast, UNCOVER heavily relies on the confined size of the sub-aperture pair to generate a tight focus. Thus, UNCOVER is fundamentally different from the AO method.

It is believed that conventional wavefront shaping is unable to generate a focus on a large, homogeneous sample, as suggested by Katz et. al [66]. Here, we showed that with constraint on the aperture size in the optimization, a tight focus can be generated without exploring non-linear feedback. However, it should be noted that if the sample is too large (say close to infinity in the extreme case), UNCOVER will not generate a tight focus as well. In this case, UNCOVER is not exempt from the discussion in [66].

PBR as a function of the initial signal

To investigate the peak-to-background ratio (PBR) that UNCOVER can achieve under different conditions, we conducted simulations by changing the initial number of photons received by the detector. In this simulation, the random phase shift (which characterizes the scattering of the wall) kept unchanged while the input light power was adapted to different levels. We also used the same target in this simulation. Apart from the signal reflected by the target, we added a background to the signal to imitate reflected photons from elsewhere, and Poisson noise was then simulated.

Figure 4.11 shows the simulation result. We iterated through the whole set of Hadamard basis for 3 times in this simulation.

We expect the PBR grows gradually as the mean initial SNR approaches 1. As we plotted the final PBR against the mean initial SNR, SNR measured using some sub-apertures can be higher than the mean SNR. That is why we see an increase that appears earlier than mean SNR of 1. The PBR increases steeply as the mean initial signal level approaches the noise level (SNR approaches 1) and then plateaus as we further increase the mean initial signal.

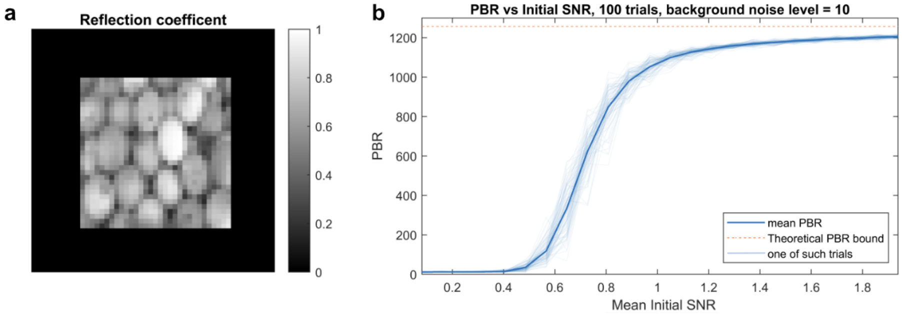

Figure 4.11: UNCOVER's peak-to-background ratio (PBR) vs mean initial SNR. **a**, reflection coefficient of the target used in this simulation. **b**, PBR as a function of the mean initial SNR (the SNR detected at the beginning of Step 2 in the optimization, averaged over all sub-apertures). In this simulation, in total $Q = 25$ sub-apertures were used, and $N^2 = 8^2 = 64$ independent modes were simulated in each sub-aperture. The color of the line in the legend representing result for one of such trials is emphasized for better illustration.

Cumulative error in updating phase offset

As aperture synthesizing is crucial for generating a sharp focus, we investigated the error in finding the relative phase offset using pairs. We require the sub-apertures in the pair to be adjacent to ensure generation of a tight focus, thus, phase error in the reference will "propagate" to the sub-aperture being updated. This means that noises accumulate when the number of sub-apertures increases. As such, studying the noise as a function of the number of sub-apertures is useful for us to understand the limits of UNCOVER.

For a better illustration, we chose to update the phase offset in a more straightforward but less sophisticated way: we updated the offset along a "snake" trajectory rather than the "spiral" one that is used in our experiments.

In this simulation, we set the mean signal detected by opening one of the sub-apertures equal to the mean value where the plateau begins in Fig. 4.11 (i.e. the mean signal value after optimization with mean initial SNR equals 1). To isolate the contribution of the noise in optimizing individual sub-apertures, we conducted the simulation in the following way:

In optimizing the phase mask for an individual sub-aperture, no noise is simulated. Poisson noise was simulated in tuning the phase offset, and we ran 5000 trials to get the distribution of the phase error. We also simulated such tuning process in absence of the noise for obtaining a ground truth of the phase offset. Phase difference in Fig. 4.12d is the difference between the ground truth and the phase offset we obtained

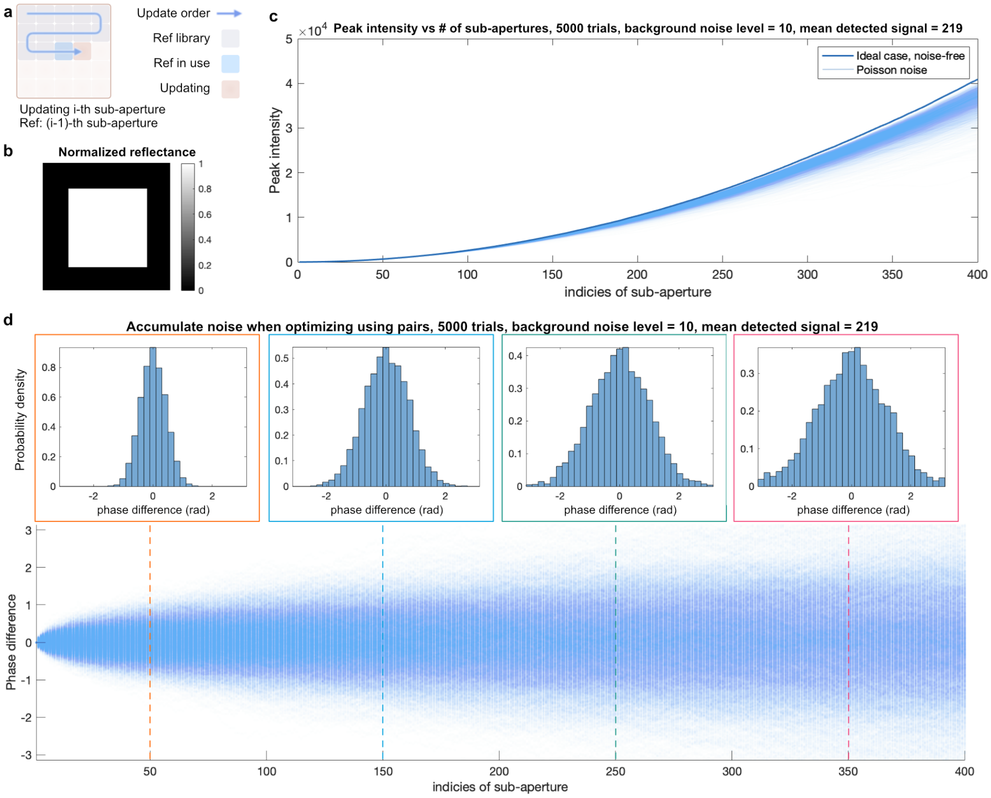

Figure 4.12: Cumulative noise in UNCOVER. **a**, tuning the phase offset along the snake line. When sub-aperture i is being updated, the one which gets updated in step $i-1$ is chosen as the reference. Thus, the noise accumulates every time we update the phase offset. **b**, the normalized reflectance of the object used in this simulation. **c**, peak intensity as a function of the number of sub-apertures. The color of the line in the legend representing result for one of such trials (denoted by "Poisson noise") is emphasized for better illustration. **d**, the phase error for each of the sub-aperture. The true phase offset is obtained in simulation when no noise is added. Figures inserted above are the histograms of the phase errors for the 50^{th}, 150^{th}, 250^{th}, and 350^{th} sub-apertures. "Mean detected signal" in **c** and **d** is the mean signal detected by opening one of the sub-apertures.

when Poisson noise is added.

The probability of the phase difference that falls in range $[-\frac{\pi}{2}, \frac{\pi}{2}]$ is around 90% up to the 240^{th} sub-aperture, as we found in our simulation. And the peak intensity when using all 400 sub-apertures is close to the ideal case for most of the trials, as Fig. 4.12c shows. If we want to keep such fidelity (probability that is greater than 90%), the distance between any sub-aperture and the center sub-aperture should be less than 240 sub-apertures. Using the the spiral order we discussed in the paper, this distant limit indicates that we can use $\sim$ 240-by-240 sub-apertures when considering the diagonal direction. Thus, the final focus can be shrunk by over 200 times compared with the one obtained by using one sub-aperture. With a higher SNR, this can be further increased.

Imaging a hollow target

In the derivation of UNCOVER, we proved that UNCOVER can generate a focus at the “center” of the target. Here, we conducted an experiment to investigate whether UNCOVER would produce a focus at the center of a hollow target, which consists of two squares as shown in Fig. 4.13a.

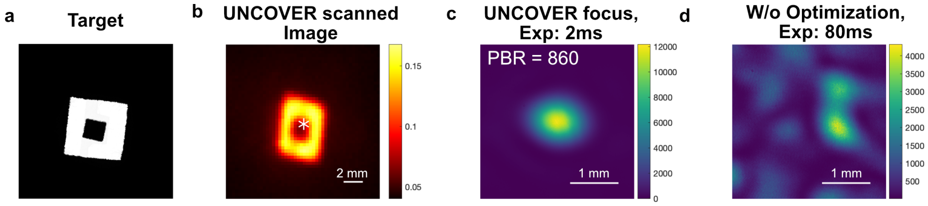

Figure 4.13: UNCOVER result of imaging a hollow object. In this experiment, $Q = 196$ sub-apertures (14×14) were used. A square of side length ≈ 0.61 mm is illuminated when using the full aperture. **a**, image of the target. **b**, UNCOVER result. The focus was initially formed at the white asterisk $*$. **c**, image of the UNCOVER focus, and **d** the speckles without performing optimization.

As shown in Fig. 4.13b, the focus located at the inner black square rather than the bright area where the reflectivity is much higher. This verifies that UNCOVER is very different from the conventional feedback based wavefront shaping. The maximizing problem UNCOVER solves is equivalent to maximizing the intensity at the center of the object, where the reflection coefficient can be as low as zero, as shown in this experiment.

UNCOVER focus quality as a function of sub-aperture size

We conclude, in the second part of Section 4.6, that UNCOVER generates a sharp focus at C_m given the criterion is satisfied. Here, we want to investigate how the shape of the final focus changes when the criterion is broken.

In this experiment, we set our parameters such that the full aperture generated by those parameters corresponded to the same area on the SLM. To be specific, the products of all the terms on the far left of Fig. 4.14 are the same, meaning that the ultimate numerical apertures for all the settings are the same. The last row of Fig.

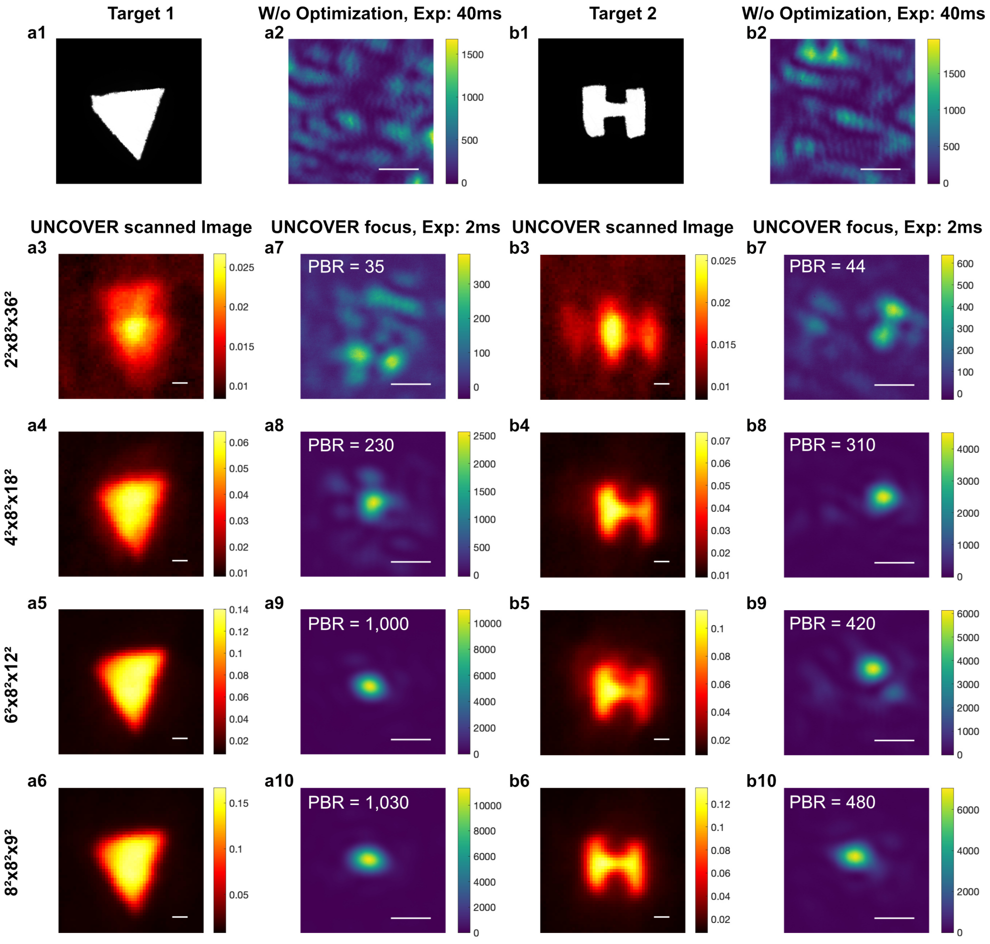

Figure 4.14: UNCOVER results when using different sub-aperture sizes. **a** and **b** are the results using target 1 and 2, respectively. The number on the far left of the figure denotes settings used for results in this row: $x^2 \times y^2 \times z^2$ means that there are x^2 sub-apertures, each sub-aperture has y^2 independent modes, and a square of z-by-z pixels on the SLM are treated as one independent mode. The side length for each sub-aperture is proportional to $y \times z$. For better illustration, we set different exposure time of the camera to capture UNCOVER foci (2 ms) and the speckle patterns without optimization (40 ms). A square of side length $\approx$ 0.35 mm is illuminated when using the full aperture. Length of the scale bars: 1 mm.

4.14 shows the results we get when the constraint of UNCOVER is satisfied. Others show the results when the constraint is broken.

When the size of the sub-aperture is too large, the “focus” deviates from an ideal one. Discrepancies between the scanned image and the real target are easy to be noticed in Fig. 4.14 a3 and b3. As we decrease the size of sub-apertures, we gradually get a single focus, as shown in Figs. 4.14 a6 and 4.14 b6. Despite of this improvement, we noticed that the PBR improvement from Fig. 4.14 b9 to Fig. 4.14 b10 is minor. This is because less photons are used in optimizing one sub-aperture, resulting a decreased SNR.

It is worth noting that the shape of the object converges to the ground truth when we gradually decrease the size of the sub-aperture until it meets the requirement in UNCOVER. This also indicates that we can use it as a metric to tell whether UNCOVER generates a single focus when it is infeasible to obtain a rough gauge of the object size.

Resolution in UNCOVER

To quantify the resolution in UNCOVER, radial profiles were extracted. The radial profile was fitted by a single Gaussian distribution. Full width at half maximum (FWHM) of the fitted curves was measured. As shown in Fig. 4.15c, the mean resolution is measured to be 0.57 mm.

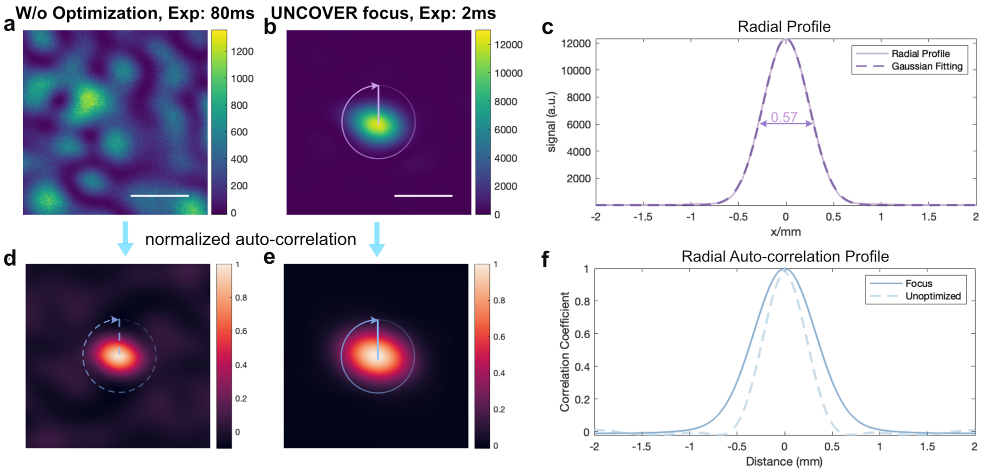

Figure 4.15: Resolution measurement of UNCOVER. This example uses the data from Fig. 5 of the main paper. **a**, speckle pattern captured before optimization. **b**, UNCOVER focus obtained after optimization. **c**, the radial profile that is obtained by averaging over a circle (one example circle is shown in **b**). Gaussian distribution is used to fit the curve. And the fitted curve is used to calculate the full width at half maximum (FWHM). **d** and **e** are the auto-correlation maps of **a** and **b**, respectively. **f**, the correlation radial profiles are averaged along the circles like those shown in **d** and **e**. Scale bar: 1 mm.

To compare the resolution of UNCOVER with the diffraction limit, auto-correlation of the unoptimized speckle pattern and the focus generated by UNCOVER were calculated. The correlation coefficient maps are shown in Fig. 4.15d and e. The

FWHM of the correlation profile of UNCOVER is in general 1.5 times wider than that of the speckle pattern. Although UNCOVER does not achieve the ideal diffraction limit, it is close to the diffraction limit, which is a significant improvement over existing ToF methods.

Collection efficiency in UNCOVER and ToF methods

ToF methods need to confine the area of detection to encode the shape of the target into the time differences of the arriving photons. This confinement in return limits the distance that ToF methods can deal with. In contrast, UNCOVER considers only the power of the detected light. Thus, almost no restriction is needed in the area on the wall being detected in UNCOVER (expect that light from the full aperture should be avoided). By imaging the detector onto the wall, we can derive that the number of photons being detected in UNCOVER is asymptotically independent of the distance between the wall and the detector. In contrast, in ToF methods, the aforementioned confinement leads to an inverse-square type drop-off in the number of photons it can detect with respect to the distance between the wall and the detector.

Here, we calculate the number of photons being detected in UNCOVER and ToF method while assuming the laser power, detector sensitivity, active area of the detector, the collecting lens, and the distances are the same in both cases.

As the distance from the wall to the the detection unit D is in general far larger than the focal length f of the lens in use, we will assume that $D \gg f$ in the following derivation. Considering the fact that the radius of the lens r is also small compared with D, we also assume $D \gg r$. When the target is far away from the wall, the photons that are scattered off by the target are approximately uniformly distributed across all points on the wall. In the following discussion, we will assume this spatial invariant intensity to be I_0. Refer to Fig. 4.16 for the definition of the different distances/lengths used in the following derivation.

In ToF method, an area of 1mm × 1mm ($A_{\text{wall}} = 1 \times 10^{-6}$ m^2) is typically detected. Let the portion of light being detected by the detector with respect to the total power shine on an unit area on the wall be P. We know that P depends on the solid angle Ω as well as the location of the unit area. As we assume that the intensity is spatially invariant, we will ignore the dependence on the location. Thus, P is a function of Ω, which will be written as $P(\Omega)$. Using Taylor expansion at $\Omega = 0$, we have:

$$P(\Omega) = a_1\Omega + a_2\Omega^2 + \cdots + a_n\Omega^n + \cdots , \tag{4.48}$$

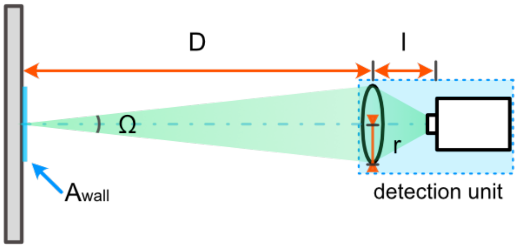

Figure 4.16: Detection scheme of UNCOVER and ToF. Ω is the acceptance angle of the collection lens (here Ω is a solid angle). We assume an area of A_{wall} is detected. The distance from the lens to the wall is D, the distance from the lens to the image sensor is l and the radius of the lens is r. The focal length of the lens is f. The angles shown in the figure are for illustration and are different from the angles in our experiments.

where $a_n (n = 1, 2, 3, \cdots)$ are constant. The constant term is dropped as the ratio is zero when $\Omega = 0$. If Ω is small, we can keep the leading term and drop $O(\Omega^2)$, which gives:

$$P(\Omega) \approx a_1 \Omega, \text{ given } \Omega \ll 1. \tag{4.49}$$

For $r \ll D$, we have the following approximation:

$$\Omega \approx \frac{\pi r^2}{D^2} \ll 1. \tag{4.50}$$

Thus, $P(\Omega)$ can be approximated as

$$P(\Omega) \approx a_1 \Omega \approx a_1 \frac{\pi r^2}{D^2}. \tag{4.51}$$

Using this approximation, the number of photons being detected per unit time in ToF methods N_{ToF} is:

$$N_{\text{ToF}} = \frac{I_0 A_{\text{wall}} P(\Omega)}{h\nu} \approx \frac{I_0 A_{\text{wall}} a_1 \frac{\pi r^2}{D^2}}{h\nu} = \frac{a_1 \pi I_0 A_{\text{wall}} r^2}{h\nu D^2}, \tag{4.52}$$

where h is the plank constant, A_{wall} the area being detected on the wall, and ν the frequency of the light.

In UNCOVER, we do not need to confine the area of detection (except the apertures should be excluded), so the area being detected is mainly dependent on the active area of the detector. Assume the area of the detector is A_{Det}, and we image the wall onto the detector. Let us first calculate l, the distance between the lens and the image plane:

$$\frac{1}{f} = \frac{1}{D} + \frac{1}{l} \Rightarrow l = \frac{fd}{f + D} \approx \frac{fd}{D} = f. \tag{4.53}$$

Thus, the magnification power M is give by:

$$M = \frac{D}{l} \approx \frac{D}{f}. \tag{4.54}$$

So, in UNCOVER, the area being detected on the wall is

$$A'_{\text{wall}} = M^2 A_{\text{Det}} = \frac{D^2}{f^2} A_{\text{Det}}. \tag{4.55}$$

Thus, we can find the number of photons being detected per unit time in UNCOVER N_{UNCOVER} is

$$N_{\text{UNCOVER}} = \frac{I'_0 A'_{\text{wall}} P(\Omega)}{h\nu} \approx \frac{I'_0 \frac{D^2}{f^2} A_{\text{Det}} a_1 \frac{\pi r^2}{D^2}}{h\nu} = \frac{a_1 \pi I'_0 A_{\text{Det}} r^2}{h\nu f^2}. \tag{4.56}$$

An important finding is that N_{ToF} is inversely proportional to D^2 while N_{UNCOVER} is independent of D. This implies that the signal that UNCOVER detected is asymptotically independent of the distance between the detector and the wall. In contrast, signal in ToF system drops as the distance increases. It is worthwhile to point out that one cannot move the detector close to the wall in many practical NLOS settings. Thus, studying the enhancement in the number of photons that can be detected in UNCOVER compared with ToF type methods is important to find the operation range improvement of UNCOVER. The ratio of number of photons being detected in UNCOVER over ToF is given by

$$\frac{N_{\text{UNCOVER}}}{N_{\text{ToF}}} = \frac{a_1 \pi I'_0 A_{\text{Det}} r^2}{h\nu f^2} \frac{h\nu D^2}{a_1 \pi I_0 A_{\text{wall}} r^2} = \frac{I'_0 A_{\text{Det}} D^2}{I_0 A_{\text{wall}} f^2}. \tag{4.57}$$

To compare the number of photons being detected in UNCOVER and ToF method, we did an experiment to verify the signal decay. The results are shown in Fig. 4.17. In the experiment, a collecting lens of focal length $f = 32$mm was used. We measured the signal starting at a distance of 0.5 away from the wall and increased the distance gradually. Each data point is the mean of 200 measurements and the length of error bar means their standard deviation. To match up with our experiment condition, the intensity for UNCOVER is decreased by 157 times ($\sim 50\pi$) as we use around 100 sub-apertures for most of our experiment. The detector used in our experiment has an active area of 3.7 mm × 13.0 mm and the area being detected in ToF is a square with side length of 1 mm.

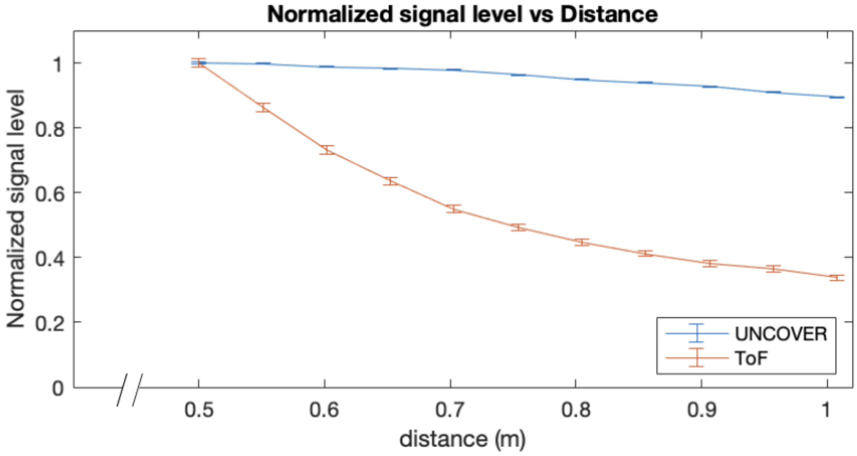

Figure 4.17: Power drop-off in UNCOVER and ToF method. The signals are normalized using their maximum value, respectively. Each point shows the mean of 200 measurements and the error bar shows the standard deviation of the 200 measurements.

Using Eq. 4.57, we can find that the ratio of signal in UNCOVER over the signal in ToF method at a distance of 0.5 m is

$$\frac{N_{\text{UNCOVER}}}{N_{\text{ToF}}} = \frac{I_0' A_{\text{Det}} D^2}{I_0 A_{\text{wall}} f^2} = \frac{1}{157} \frac{3.7 \times 13}{1 \times 1} \frac{0.5^2}{0.032^2} \approx 75. \tag{4.58}$$

In the actual experiment, we find that the ratio is approximately 87, which is slightly higher than the value based on our calculation. This is mainly due the fact that the scatted photon returning to the wall is not evenly distributed. In UNCOVER, a much larger area is being detected, which covers areas with higher intensities. As shown in Fig. 4.17, signal in ToF decreases much faster than UNCOVER. This shows that UNCOVER may potentially provide a larger operation range than the existing ToF methods. We note that the signals in UNCOVER exhibit a gentle decrease when the distance from the wall to the detector increases. This might also be caused by the uneven distribution of the scattered light, which violates the assumption that the intensity is spatially invariant.

In addition to the direct improvement in the detection efficiency, we want to point out that conducting feedback based wavefront shaping directly (without using pairs) also benefits in UNCOVER. Although applying feedback based wavefront shaping directly without using pairs is not guaranteed to, and probably will not, provide a focus, it delivers more light onto the target. With this pre-optimization, operation range of UNCOVER can be further increased.

BIBLIOGRAPHY

1. Goodman, J. W. *Introduction to Fourier Optics* ISBN: 978-0-9747077-2-3 (Roberts and Company Publishers, 2005).
2. Fienup, J. R. Phase retrieval algorithms: A comparison. *Applied Optics* **21.** Publisher: Optica Publishing Group, 2758–2769. ISSN: 2155-3165. `https://opg.optica.org/ao/abstract.cfm?uri=ao-21-15-2758` (Aug. 1, 1982).
3. Zheng, G., Horstmeyer, R. & Yang, C. Wide-field, high-resolution Fourier ptychographic microscopy. *Nature Photonics* **7,** 739–745. ISSN: 1749-4893. `https://www.nature.com/articles/nphoton.2013.187` (2013).
4. Zheng, G., Shen, C., Jiang, S., Song, P. & Yang, C. Concept, implementations and applications of Fourier ptychography. *Nature Reviews Physics* **3.** Number: 3 Publisher: Nature Publishing Group, 207–223. ISSN: 2522-5820. `https://www.nature.com/articles/s42254-021-00280-y` (2021).
5. Ou, X., Zheng, G. & Yang, C. Embedded pupil function recovery for Fourier ptychographic microscopy. *Optics Express* **22.** Publisher: Optica Publishing Group, 4960–4972. ISSN: 1094-4087. `https://opg.optica.org/oe/abstract.cfm?uri=oe-22-5-4960` (2014).
6. Konda, P. C. *et al.* Fourier ptychography: Current applications and future promises. *Opt. Express* **28.** Publisher: Optica Publishing Group, 9603–9630. `http://opg.optica.org/oe/abstract.cfm?URI=oe-28-7-9603` (Mar. 2020).
7. Tian, L., Li, X., Ramchandran, K. & Waller, L. Multiplexed coded illumination for Fourier Ptychography with an LED array microscope. *Biomedical Optics Express* **5.** Publisher: Optica Publishing Group, 2376–2389. ISSN: 2156-7085. `https://opg.optica.org/boe/abstract.cfm?uri=boe-5-7-2376` (2014).
8. Tian, L. *et al.* Computational illumination for high-speed in vitro Fourier ptychographic microscopy. *Optica* **2,** 904–911. ISSN: 2334-2536. `https://www.osapublishing.org/optica/abstract.cfm?uri=optica-2-10-904` (2015).
9. Yeh, L.-H. *et al.* Experimental robustness of Fourier ptychography phase retrieval algorithms. *Optics Express* **23.** Publisher: Optica Publishing Group, 33214–33240. ISSN: 1094-4087. `https://opg.optica.org/oe/abstract.cfm?uri=oe-23-26-33214` (2015).
10. Candès, E. J., Li, X. & Soltanolkotabi, M. Phase Retrieval via Wirtinger Flow: Theory and Algorithms. *IEEE Transactions on Information Theory* **61.**

Conference Name: IEEE Transactions on Information Theory, 1985–2007. ISSN: 1557-9654 (Apr. 2015).

11. Bian, L. *et al.* Fourier ptychographic reconstruction using Wirtinger flow optimization. *Optics Express* **23.** Publisher: Optica Publishing Group, 4856–4866. ISSN: 1094-4087. `https://opg.optica.org/oe/abstract.cfm?uri=oe-23-4-4856` (Feb. 23, 2015).

12. Li, E., Sherwin, S., Gunjala, G. & Waller, L. Exceeding the limits of algorithmic self-calibrated aberration recovery in Fourier ptychography. *Optics Continuum* **2.** Publisher: Optica Publishing Group, 119–130. ISSN: 2770-0208. `https://opg.optica.org/optcon/abstract.cfm?uri=optcon-2-1-119` (Jan. 15, 2023).

13. Nguyen, T., Xue, Y., Li, Y., Tian, L. & Nehmetallah, G. Deep learning approach for Fourier ptychography microscopy. *Optics Express* **26.** Publisher: Optica Publishing Group, 26470–26484. ISSN: 1094-4087. `https://opg.optica.org/oe/abstract.cfm?uri=oe-26-20-26470` (Oct. 1, 2018).

14. Jiang, S., Guo, K., Liao, J. & Zheng, G. Solving Fourier ptychographic imaging problems via neural network modeling and TensorFlow. *Biomedical Optics Express* **9.** Publisher: Optica Publishing Group, 3306–3319. ISSN: 2156-7085. `https://opg.optica.org/boe/abstract.cfm?uri=boe-9-7-3306` (July 1, 2018).

15. Sun, M. *et al.* Neural network model combined with pupil recovery for Fourier ptychographic microscopy. *Optics Express* **27.** Publisher: Optica Publishing Group, 24161–24174. ISSN: 1094-4087. `https://opg.optica.org/oe/abstract.cfm?uri=oe-27-17-24161` (2023) (Aug. 19, 2019).

16. Zhou, K. C. & Horstmeyer, R. Diffraction tomography with a deep image prior. *Optics Express* **28.** Publisher: Optica Publishing Group, 12872–12896. ISSN: 1094-4087. `https://opg.optica.org/oe/abstract.cfm?uri=oe-28-9-12872` (Apr. 27, 2020).

17. Baek, Y. & Park, Y. Intensity-based holographic imaging via space-domain Kramers–Kronig relations. *Nature Photonics* **15.** Number: 5 Publisher: Nature Publishing Group, 354–360. ISSN: 1749-4893. `https://www.nature.com/articles/s41566-021-00760-8` (2021).

18. Baek, Y., Lee, K., Shin, S. & Park, Y. Kramers–Kronig holographic imaging for high-space-bandwidth product. *Optica* **6.** Publisher: Optica Publishing Group, 45–51. ISSN: 2334-2536. `https://opg.optica.org/optica/abstract.cfm?uri=optica-6-1-45` (2019).

19. Shen, C., Liang, M., Pan, A. & Yang, C. Non-iterative complex wave-field reconstruction based on Kramers–Kronig relations. *Photonics Research* **9.** Publisher: Optica Publishing Group, 1003–1012. ISSN: 2327-9125. `https://opg.optica.org/prj/abstract.cfm?uri=prj-9-6-1003` (2021).

20. Havlicek, J., Havlicek, J. & Bovik, A. *The analytic image* in *Proceedings of International Conference on Image Processing* Proceedings of International Conference on Image Processing. **2** (1997), 446–449 vol.2.

21. Ghiglia, D. C. & Romero, L. A. Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. *JOSA A* **11.** Publisher: Optica Publishing Group, 107–117. ISSN: 1520-8532. `https://opg.optica.org/josaa/abstract.cfm?uri=josaa-11-1-107` (1994).

22. Eckert, R., Phillips, Z. F. & Waller, L. Efficient illumination angle self-calibration in Fourier ptychography. *Applied Optics* **57.** Publisher: Optica Publishing Group, 5434–5442. ISSN: 2155-3165. `https://opg.optica.org/ao/abstract.cfm?uri=ao-57-19-5434` (2018).

23. Wang, L. V. & Hu, S. Photoacoustic tomography: In vivo imaging from organelles to organs. *Science* **335.** Publisher: American Association for the Advancement of Science, 1458–1462. `https://www.science.org/doi/10.1126/science.1216210` (Mar. 23, 2012).

24. Elson, D. S., Li, R., Dunsby, C., Eckersley, R. & Tang, M.-X. Ultrasound-mediated optical tomography: A review of current methods. *Interface Focus* **1.** Publisher: Royal Society, 632–648. `https://royalsocietypublishing.org/doi/10.1098/rsfs.2011.0021` (June 2, 2011).

25. Cuche, E., Marquet, P. & Depeursinge, C. Spatial filtering for zero-order and twin-image elimination in digital off-axis holography. *Applied Optics* **39.** Publisher: Optica Publishing Group, 4070–4075. ISSN: 2155-3165. `https://opg.optica.org/ao/abstract.cfm?uri=ao-39-23-4070` (2023) (Aug. 10, 2000).

26. Durduran, T., Choe, R., Baker, W. B. & Yodh, A. G. Diffuse optics for tissue monitoring and tomography. *Reports on Progress in Physics* **73,** 076701. ISSN: 0034-4885. `https://dx.doi.org/10.1088/0034-4885/73/7/076701` (2023) (June 2010).

27. Qureshi, M. M. *et al.* In vivo study of optical speckle decorrelation time across depths in the mouse brain. *Biomedical Optics Express* **8.** Publisher: Optica Publishing Group, 4855–4864. ISSN: 2156-7085. `https://opg.optica.org/boe/abstract.cfm?uri=boe-8-11-4855` (2023) (Nov. 1, 2017).

28. Liu, Y. *et al.* Optical focusing deep inside dynamic scattering media with near-infrared time-reversed ultrasonically encoded (TRUE) light. *Nature Communications* **6.** Number: 1 Publisher: Nature Publishing Group, 5904. ISSN: 2041-1723. `https://www.nature.com/articles/ncomms6904` (Jan. 5, 2015).

29. Lévêque, S., Boccara, A. C., Lebec, M. & Saint-Jalmes, H. Ultrasonic tagging of photon paths in scattering media: parallel speckle modulation processing.

Optics Letters **24.** Publisher: Optica Publishing Group, 181–183. ISSN: 1539-4794. https://opg.optica.org/ol/abstract.cfm?uri=ol-24-3-181 (2023) (Feb. 1, 1999).

30. Li, J., Ku, G. & Wang, L. V. Ultrasound-modulated optical tomography of biological tissue by use of contrast of laser speckles. *Applied Optics* **41.** Publisher: Optica Publishing Group, 6030–6035. ISSN: 2155-3165. https://opg.optica.org/ao/abstract.cfm?uri=ao-41-28-6030 (2023) (Oct. 1, 2002).
31. Liu, Y., Shen, Y., Ma, C., Shi, J. & Wang, L. V. Lock-in camera based heterodyne holography for ultrasound-modulated optical tomography inside dynamic scattering media. *Applied Physics Letters* **108,** 231106. ISSN: 0003-6951. https://doi.org/10.1063/1.4953630 (2023) (June 8, 2016).
32. Barjean, K. *et al.* Fourier transform acousto-optic imaging with a custom-designed CMOS smart-pixels array. *Optics Letters* **40.** Publisher: Optica Publishing Group, 705–708. ISSN: 1539-4794. https://opg.optica.org/ol/abstract.cfm?uri=ol-40-5-705 (2023) (Mar. 1, 2015).
33. Liu, Y., Ma, C., Shen, Y. & Wang, L. V. Bit-efficient, sub-millisecond wavefront measurement using a lock-in camera for time-reversal based optical focusing inside scattering media. *Optics Letters* **41.** Publisher: Optica Publishing Group, 1321–1324. ISSN: 1539-4794. https://opg.optica.org/ol/abstract.cfm?uri=ol-41-7-1321 (Apr. 1, 2016).
34. Gross, M., Goy, P. & Al-Koussa, M. Shot-noise detection of ultrasound-tagged photons in ultrasound-modulated optical imaging. *Optics Letters* **28.** Publisher: Optica Publishing Group, 2482–2484. ISSN: 1539-4794. https://opg.optica.org/ol/abstract.cfm?uri=ol-28-24-2482 (2023) (Dec. 15, 2003).
35. Ruan, H. *et al.* Deep tissue optical focusing and optogenetic modulation with time-reversed ultrasonically encoded light. *Science Advances* **3,** eaao5520. ISSN: 2375-2548. https://advances.sciencemag.org/content/3/12/eaao5520 (2019) (Dec. 1, 2017).
36. Aubry, J.-F., Tanter, M., Pernot, M., Thomas, J.-L. & Fink, M. Experimental demonstration of noninvasive transskull adaptive focusing based on prior computed tomography scans. *The Journal of the Acoustical Society of America* **113,** 84–93. ISSN: 0001-4966. https://doi.org/10.1121/1.1529663 (2023) (Jan. 8, 2003).
37. Huang, C. *et al.* Aberration correction for transcranial photoacoustic tomography of primates employing adjunct image data. *Journal of Biomedical Optics* **17.** Publisher: SPIE, 066016. ISSN: 1083-3668, 1560-2281. https://www.spiedigitallibrary.org/journals/journal-of-biomedical-optics/volume-17/issue-6/066016/Aberration-correction-

`for-transcranial-photoacoustic-tomography-of-primates-employing-adjunct/10.1117/1.JBO.17.6.066016.full` (June 2012).

38. Chauvet, D. *et al.* Targeting accuracy of transcranial magnetic resonance–guided high-intensity focused ultrasound brain therapy: A fresh cadaver model: Laboratory investigation. *Journal of Neurosurgery* **118.** Publisher: American Association of Neurological Surgeons Section: Journal of Neurosurgery, 1046–1052. ISSN: 1933-0693, 0022-3085. `https://thejns.org/view/journals/j-neurosurg/118/5/article-p1046.xml` (May 1, 2013).

39. Ma, C., Zhou, F., Liu, Y. & Wang, L. V. Single-exposure optical focusing inside scattering media using binarized time-reversed adapted perturbation. *Optica* **2.** Publisher: Optica Publishing Group, 869–876. ISSN: 2334-2536. `https://opg.optica.org/optica/abstract.cfm?uri=optica-2-10-869` (Oct. 20, 2015).

40. Faccio, D., Velten, A. & Wetzstein, G. Non-line-of-sight imaging. *Nature Reviews Physics* **2,** 318–327. ISSN: 2522-5820. `https://www.nature.com/articles/s42254-020-0174-8` (2020).

41. Bertolotti, J. *et al.* Non-invasive imaging through opaque scattering layers. *Nature* **491,** 232–234. ISSN: 1476-4687. `https://www.nature.com/articles/nature11578` (2012).

42. Katz, O., Heidmann, P., Fink, M. & Gigan, S. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nature Photonics* **8,** 784–790. ISSN: 1749-4893. `https://www.nature.com/articles/nphoton.2014.189` (2014).

43. Velten, A. *et al.* Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Communications* **3,** 745. ISSN: 2041-1723. `https://www.nature.com/articles/ncomms1747` (2012).

44. O'Toole, M., Lindell, D. B. & Wetzstein, G. Confocal non-line-of-sight imaging based on the light-cone transform. *Nature* **555,** 338–341. ISSN: 1476-4687. `https://www.nature.com/articles/nature25489` (2018).

45. Xin, S. *et al. A Theory of Fermat Paths for Non-Line-Of-Sight Shape Reconstruction* in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), 6793–6802.

46. Liu, X. *et al.* Non-line-of-sight imaging using phasor-field virtual wave optics. *Nature* **572,** 620–623. ISSN: 1476-4687. `https://www.nature.com/articles/s41586-019-1461-3` (2019).

47. Reza, S. A., Manna, M. L., Bauer, S., Velten, A. & Velten, A. Phasor field waves: A Huygens-like light transport model for non-line-of-sight imaging applications. *Optics Express* **27,** 29380–29400. ISSN: 1094-4087. `https://www.osapublishing.org/oe/abstract.cfm?uri=oe-27-20-29380` (2019).

48. Gupta, M., Nayar, S. K., Hullin, M. B. & Martin, J. Phasor imaging: A generalization of correlation-based time-of-flight Imaging. *ACM Transactions on Graphics* **34,** 156:1–156:18. ISSN: 0730-0301. `https://doi.org/10.1145/2735702` (2015).

49. Saunders, C., Murray-Bruce, J. & Goyal, V. K. Computational periscopy with an ordinary digital camera. *Nature* **565,** 472. ISSN: 1476-4687. `https://www.nature.com/articles/s41586-018-0868-6` (2019).

50. Vellekoop, I. M. & Mosk, A. P. Focusing coherent light through opaque strongly scattering media. *Optics Letters* **32,** 2309–2311. ISSN: 1539-4794. `https://www.osapublishing.org/ol/abstract.cfm?uri=ol-32-16-2309` (2007).

51. Yaqoob, Z., Psaltis, D., Feld, M. S. & Yang, C. Optical phase conjugation for turbidity suppression in biological samples. *Nature Photonics* **2,** 110–115. ISSN: 1749-4893. `https://www.nature.com/articles/nphoton.2007.297` (2008).

52. Popoff, S., Lerosey, G., Fink, M., Boccara, A. C. & Gigan, S. Image transmission through an opaque material. *Nature Communications* **1,** 81. ISSN: 2041-1723. `https://www.nature.com/articles/ncomms1078` (2010).

53. Popoff, S. M., Lerosey, G., Fink, M., Boccara, A. C. & Gigan, S. Controlling light through optical disordered media: Transmission matrix approach. *New Journal of Physics* **13,** 123021. ISSN: 1367-2630. `https://doi.org/10.1088%2F1367-2630%2F13%2F12%2F123021` (2011).

54. Mounaix, M. *et al.* Spatiotemporal coherent control of light through a multiple scattering medium with the multispectral transmission matrix. *Physical Review Letters* **116,** 253901. `https://link.aps.org/doi/10.1103/PhysRevLett.116.253901` (2016).

55. Conkey, D. B., Caravaca-Aguirre, A. M. & Piestun, R. High-speed scattering medium characterization with application to focusing light through turbid media. *Optics Express* **20,** 1733–1740. ISSN: 1094-4087. `https://www.osapublishing.org/oe/abstract.cfm?uri=oe-20-2-1733` (2012).

56. Vellekoop, I. M. Feedback-based wavefront shaping. *Optics Express* **23,** 12189–12206. ISSN: 1094-4087. `https://www.osapublishing.org/oe/abstract.cfm?uri=oe-23-9-12189` (2015).

57. Mosk, A. P., Lagendijk, A., Lerosey, G. & Fink, M. Controlling waves in space and time for imaging and focusing in complex media. *Nature Photonics* **6,** 283–292. ISSN: 1749-4893. `https://www.nature.com/articles/nphoton.2012.88` (2012).

58. Xu, X., Liu, H. & Wang, L. V. Time-reversed ultrasonically encoded optical focusing into scattering media. *Nature Photonics* **5,** 154–157. ISSN: 1749-4893. `https://www.nature.com/articles/nphoton.2010.306` (2011).

59. Horstmeyer, R., Ruan, H. & Yang, C. Guidestar-assisted wavefront-shaping methods for focusing light into biological tissue. *Nature Photonics* **9,** 563–571. ISSN: 1749-4893. `https://www.nature.com/articles/nphoton.2015.140` (2015).

60. Ruan, H., Jang, M., Judkewitz, B. & Yang, C. Iterative time-reversed ultrasonically encoded light focusing in backscattering mode. *Scientific Reports* **4,** 1–7. ISSN: 2045-2322. `https://www.nature.com/articles/srep07156` (2014).

61. Katz, O., Ramaz, F., Gigan, S. & Fink, M. Controlling light in complex media beyond the acoustic diffraction-limit using the acousto-optic transmission matrix. *Nature Communications* **10,** 1–10. ISSN: 2041-1723. `https://www.nature.com/articles/s41467-019-08583-6` (2019).

62. Judkewitz, B., Wang, Y. M., Horstmeyer, R., Mathy, A. & Yang, C. Speckle-scale focusing in the diffusive regime with time reversal of variance-encoded light (TROVE). *Nature Photonics* **7,** 300–305. ISSN: 1749-4893. `https://www.nature.com/articles/nphoton.2013.31` (2013).

63. Conkey, D. B. *et al.* Super-resolution photoacoustic imaging through a scattering wall. *Nature Communications* **6,** 7902. ISSN: 2041-1723. `https://doi.org/10.1038/ncomms8902` (2015).

64. Katz, O., Small, E. & Silberberg, Y. Looking around corners and through thin turbid layers in real time with scattered incoherent light. *Nature Photonics* **6,** 549–553. ISSN: 1749-4885, 1749-4893. `http://www.nature.com/articles/nphoton.2012.150` (2012).

65. Starshynov, I., Ghafur, O., Fitches, J. & Faccio, D. Coherent Control of Light for Non-Line-of-Sight Imaging. *Physical Review Applied* **12,** 064045. `https://link.aps.org/doi/10.1103/PhysRevApplied.12.064045` (2019).

66. Katz, O., Small, E., Guan, Y. & Silberberg, Y. Noninvasive nonlinear focusing and imaging through strongly scattering turbid layers. *Optica* **1,** 170–174. ISSN: 2334-2536. `https://www.osapublishing.org/optica/abstract.cfm?uri=optica-1-3-170` (2014).

67. Osnabrugge, G., Horstmeyer, R., Papadopoulos, I. N., Judkewitz, B. & Vellekoop, I. M. Generalized optical memory effect. *Optica* **4,** 886–892. ISSN: 2334-2536. `https://www.osapublishing.org/optica/abstract.cfm?uri=optica-4-8-886` (2017).

68. Freund, I., Rosenbluh, M. & Feng, S. Memory effects in propagation of optical waves through disordered media. *Physical Review Letters* **61,** 2328–2331. `https://link.aps.org/doi/10.1103/PhysRevLett.61.2328` (1988).

69. Jang, J. *et al.* Complex wavefront shaping for optimal depth-selective focusing in optical coherence tomography. *Optics Express* **21,** 2890–2902. ISSN: 1094-4087. `https://www.osapublishing.org/oe/abstract.cfm?uri=oe-21-3-2890` (2013).

70. Ji, N., Milkie, D. E. & Betzig, E. Adaptive optics via pupil segmentation for high-resolution imaging in biological tissues. *Nature Methods* **7,** 141–147. ISSN: 1548-7105. `https://www.nature.com/articles/nmeth.1411` (2010).

Appendix A

NOISE IN OFFAXIS HOLOGRAPHY

We will calculate the signal power (intensity of the signal) using offaxis holography. To do this, we will need to know the noise distribution after the process in offaxis. Thus, we will first find the noise distribution after we perform Fourier transform to the acquired image. Then, we will use this information to deduce the noise distribution after cropping and inverse Fourier transform. SNR here refers to SNR per pixel after we perform offaxis (otherwise it is clarified).

A.1 Noise distribution in the Fourier transform

For the intensity captured by one of the pixel at position $\mathbf{x}_t$ on a camera in offaxis holography, we have

$$\text{Offaxis}(\mathbf{x}_t) = R(\mathbf{x}_t) + T(\mathbf{x}_t) + 2\sqrt{RT(\mathbf{x}_t)}\cos(\mathbf{k}_0 \cdot \mathbf{x}_t - \phi) + N(\mathbf{x}_t), \tag{A.1}$$

where T and R are the signal and reference (reference is assumed to be a plane wave $R(\mathbf{x}_t) = R$), respectively. $\mathbf{k}_0$ is related to the tilt between the reference arm and signal arm, ϕ is the phase of the E-field of the signal arm. When the power of the reference is large, the noise can be approximated as a Gaussian noise with mean 0 and variance R

$$N(\mathbf{x}_t) \sim \mathcal{N}(0, \sigma^2 = R). \tag{A.2}$$

Note that the sequence $\{N(\mathbf{x}_t)\}$ is i.i.d., so they are jointly Gaussian. The DFT of the noise term can be written as (N_P is the number of pixels):

$$N(\mathbf{k}) = \sum_{t=1}^{N_P} N(\mathbf{x}_t) e^{-i\mathbf{k}\cdot\mathbf{x}_t} = \sum_{t=1}^{N_P} N(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t) - i\sum_{t=1}^{N_P} N(\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t). \tag{A.3}$$

We can deduce that the real and imaginary part are both Gaussian from the fact $\{N(\mathbf{x}_t)\}$ is jointly Gaussian. To find out if the real and imaginary part are independent, let us calculate the covariance between them, this gives

$$\begin{aligned}\mathbb{E}\Bigg[-\Bigg(\sum_{t=1}^{N_P} N(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\Bigg)\Bigg(\sum_{t}^{N_P} N(\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t)\Bigg)\Bigg]& \\ &= -\mathbb{E}\Bigg[\sum_{t=1}^{N_P} N^2(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t)\Bigg] \\ &= -\sum_{t=1}^{N_P}\mathbb{E}\left[N^2(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t)\right] \\ &= -\sigma^2\sum_{t=1}^{N_P}\cos(\mathbf{k}\cdot\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t) \\ &= 0. \end{aligned} \tag{A.4}$$

The last equality follows the fact that sin and cos are elements of Fourier basis. So the covariance of the two Gaussian R.V. is zero, which means the real and imaginary part are independent. For different $\mathbf{k}$, $\sin(\mathbf{k}_1)$, $\cos(\mathbf{k}_1)$, $\sin(\mathbf{k}_2)$ and $\cos(\mathbf{k}_2)$ ($\mathbf{k}_1 \notin \{\mathbf{k}_2, -\mathbf{k}_2\}$) are also orthogonal. Thus, $N(\mathbf{k}_1)$ and $N(\mathbf{k}_2)$ are also independent ($\mathbf{k}_1 \notin \{\mathbf{k}_2, -\mathbf{k}_2\}$). (Eq. A.4 is actually the imaginary part of the Fourier transform of the constant $\frac{\sigma^2}{2}$ evaluated at $2\mathbf{k}$, which is zero.)

Let us then calculate the variance of the centered real and imaginary part

$$\begin{aligned}\mathrm{Var}\Bigg[\sum_{t=1}^{N_P} N(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\Bigg]& \\ &= \mathbb{E}\left[\left(\sum_{t=1}^{N_P} N(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\right)^2\right] - \left[\mathbb{E}\left(\sum_{t=1}^{N_P} N(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\right)\right]^2 \\ &= \sum_{t=1}^{N_P}\mathbb{E}\left[N^2(\mathbf{x}_t)\cos^2(\mathbf{k}\cdot\mathbf{x}_t)\right] \\ &= \sigma^2\sum_{t=1}^{N_P}\cos^2(\mathbf{k}\cdot\mathbf{x}_t) = \frac{N_P}{2}\sigma^2. \end{aligned} \tag{A.5}$$

Similarly, the variance of the imaginary part is also $\frac{N_P}{2}\sigma^2$. So the modulus of the complex number $|N(\mathbf{k})| \sim \mathrm{Rayleigh}(\sigma\sqrt{\frac{N_P}{2}})$, the angle is uniformly distributed.

The following figure (Fig. A.1) shows the comparison between the distribution of the noise in the Fourier Transform and the theoretical distribution, which matches pretty well. The correlation between the imaginary and real part is -6×10^{-4}.

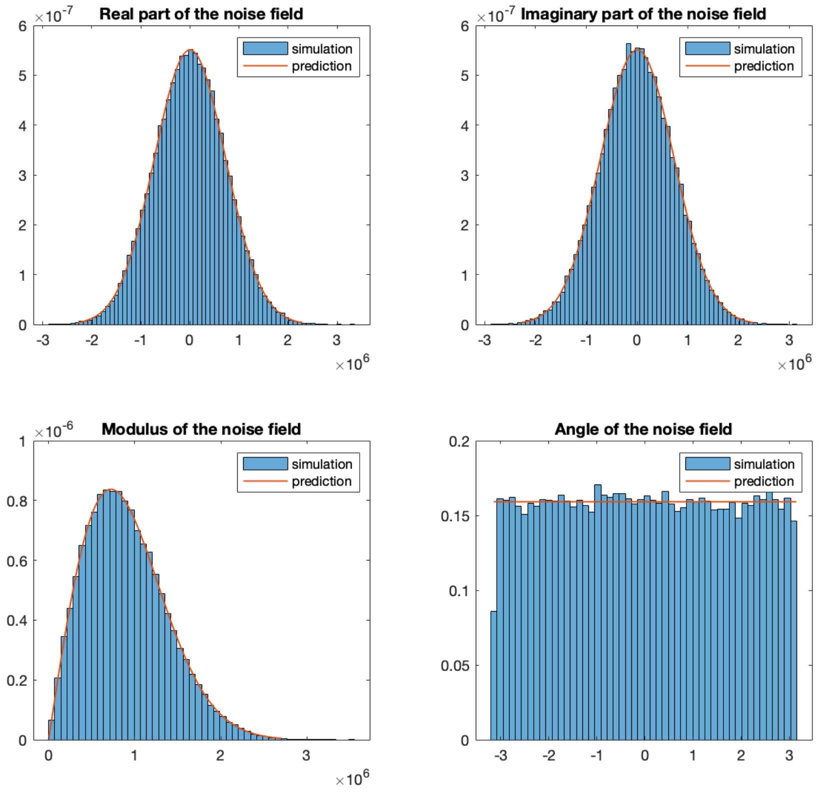

Figure A.1: Noise distribution in offaxis holography (spatial frequency space). Here, we simulated the noise using the forward model and then conducted the Fourier transform. The noise was extracted and then plotted against our theoretical prediction.

A.2 The cropping process in offaxis and its influence

From Section A.1 we know that the real and imaginary part of the FT is i.i.d. Gaussian with mean 0 and variance $\sigma'^2 = \frac{N_P}{2}\sigma^2$. Let us write $N(\mathbf{k}_j)$ in the following way:

$$N(\mathbf{k}_j) = \xi_R(\mathbf{k}_j) + i\xi_I(\mathbf{k}_j), \tag{A.6}$$

where ξ_R, ξ_I are i.i.d. $\mathcal{N}(0, \sigma'^2)$. Let us consider only half of the 2D k-space $A_h := \{\mathbf{k} = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} | k_1 < 0\}$. Note that $\forall \mathbf{k}_1 \mathbf{k}_2 \in A_h$, $N(\mathbf{k}_1) \perp N(\mathbf{k}_2)$. Suppose we cropped out a rectangle $C \subset A_h$. Let N_C denote the number of $\mathbf{k} \in C$. Consider the inverse Fourier transform on the rectangle, the signal we will end up getting is

$$\begin{aligned}\hat{N}(\mathbf{r}_m) &= \frac{1}{N_C}\sum_{j=1}^{N_C} N(\mathbf{k}_j)e^{i(\mathbf{k}_j-\mathbf{k}_0)\cdot\mathbf{r}_m} = \frac{1}{N_C}\sum_{j=1}^{N_C} N(\mathbf{k}_j)e^{i\mathbf{k}'_j\cdot\mathbf{r}_m}\\ &= \frac{1}{N_C}\sum_{j=1}^{N_C}\left[\cos(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_R(\mathbf{k}_j) - \sin(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_I(\mathbf{k}_j)\right]\\ &\quad + i\frac{1}{N_C}\sum_{j=1}^{N_C}\left[\cos(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_I(\mathbf{k}_j) + \sin(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_R(\mathbf{k}_j)\right],\end{aligned} \tag{A.7}$$

where $\mathbf{k}'_j := \mathbf{k}_j - \mathbf{k}'_0$, and $\mathbf{k}'_0$ is the center of the rectangle C. We can also conclude that the real and imaginary part of $\hat{N}(\mathbf{r}_m)$ are Gaussian. (It follows the same logic as in Section A.1.)

Let us repeat the process we went through in Section A.1. The covariance of the real and imaginary part is

$$\begin{aligned}&\mathbb{E}\left[\mathfrak{R}\left(\hat{N}(\mathbf{r}_m)\right)\mathfrak{I}\left(\hat{N}(\mathbf{r}_m)\right)\right]\\ &= \frac{1}{N_C^2}\sum_{j=1}^{N_C}\mathbb{E}\left[\cos(\mathbf{k}'_j\cdot\mathbf{r}_m)\sin(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_R^2(\mathbf{k}_j)\right]\\ &\qquad - \frac{1}{N_C^2}\sum_{j=1}^{N_C}\mathbb{E}\left[\cos(\mathbf{k}'_j\cdot\mathbf{r}_m)\sin(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_I^2(\mathbf{k}_j)\right]\\ &= \frac{\sigma'^2}{N_C^2}\sum_{j=1}^{N_C}\cos(\mathbf{k}'_j\cdot\mathbf{r}_m)\sin(\mathbf{k}'_j\cdot\mathbf{r}_m) - \frac{\sigma'^2}{N_C^2}\sum_{j=1}^{N_C}\cos(\mathbf{k}'_j\cdot\mathbf{r}_m)\sin(\mathbf{k}'_j\cdot\mathbf{r}_m)\\ &= 0,\end{aligned} \tag{A.8}$$

where $\mathfrak{R}(\cdot)$ denotes the real part of a complex number and $\mathfrak{I}(\cdot)$ denotes the imaginary part of a complex number. This indicates that the real and imaginary part of $\hat{N}(\mathbf{r}_m)$ are independent. Also, for different m and m', $\sin(\mathbf{r}_m\cdot\mathbf{k}'_j)$, $\cos(\mathbf{r}_m\cdot\mathbf{k}'_j)$, $\sin(\mathbf{r}_{m'}\cdot\mathbf{k}'_j)$ and $\cos(\mathbf{r}_{m'}\cdot\mathbf{k}'_j)$ are orthogonal. So $\hat{N}(\mathbf{r}_m)$ and $\hat{N}(\mathbf{r}'_m)$ are independent.

Note that if we choose another cropping method, the noise term may become dependent. A simple way to see this result is observing that if some of the term are dropped out, the number of degree of freedom in k-space is reduced. As the total number of freedom will not change in the inverse Fourier transform, $\{\hat{N}(\mathbf{r}_m)\}$ cannot be independent.

In the following analysis, we will just consider we cropped out the rectangle C. The variance of the real part of $\hat{N}(\mathbf{r}_m)$ is

$$\begin{aligned}\mathrm{Var}\left[\Re\left(\hat{N}(\mathbf{r}_m)\right)\right] &= \mathbb{E}\left[\left(\frac{1}{N_C}\sum_{j=1}^{N_C}\left[\cos(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_R(\mathbf{k}_j) - \sin(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_I(\mathbf{k}_j)\right]\right)^2\right] \\ &= \frac{1}{N_C^2}\left[\sum_{j=1}^{N_C}\mathbb{E}\left[\cos^2(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_R^2(\mathbf{k}_j)\right] + \sum_{j=1}^{N_C}\mathbb{E}\left[\sin^2(\mathbf{k}'_j\cdot\mathbf{r}_m)\xi_I^2(\mathbf{k}_j)\right]\right] \\ &= \frac{1}{N_C^2}\left[\sigma'^2\sum_{j=1}^{N_C}\left(\cos^2(\mathbf{k}'_j\cdot\mathbf{r}_m) + \sin^2(\mathbf{k}'_j\cdot\mathbf{r}_m)\right)\right] \\ &= \frac{\sigma'^2}{N_C} = \frac{\sigma^2 N_P}{2N_C}. \end{aligned} \tag{A.9}$$

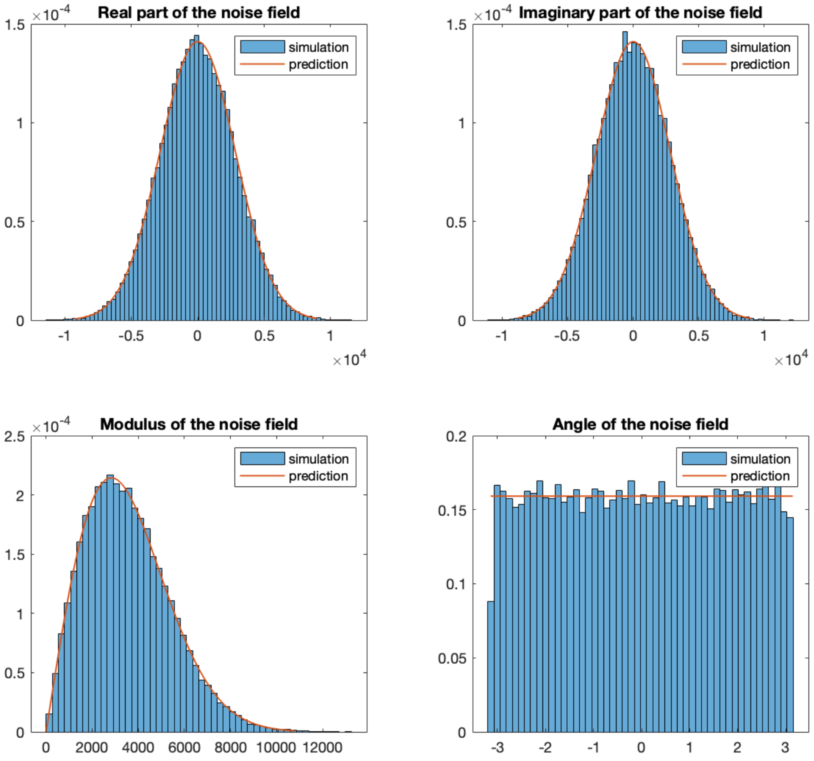

Figure A.2: Noise distribution in offaxis holography (after reconstruction). In this simulation, we performed the entire reconstruction process in offaxis holography. The cropped spectrum was shifted to the zero spatial frequency and then transformed back to real space. The histogram of the simulated noise in the reconstruction was then plotted against the prediction.

For the imaginary part, we will also have $\mathrm{Var}\left[\Im\left(\hat{N}(\mathbf{r}_m)\right)\right] = \frac{\sigma'^2}{N_C}$. If we write the noise in polar form, we will have

$$\hat{N}(\mathbf{r}_m) = A(\mathbf{r}_m)e^{i\alpha(\mathbf{r}_m)}, \tag{A.10}$$

where $A(\mathbf{r}_m) \sim \text{Rayleigh}(\frac{\sigma'}{\sqrt{N_C}})$, and $\alpha(\mathbf{r}_m) \sim \text{Unif}[0, 2\pi]$. Figure A.2 shows the comparison between theory and simulation. Similarly, the average absolute correlation between two different $\hat{N}(\mathbf{r}_m)$ is 0.003. And this simulation result also supports that $\{\hat{N}(\mathbf{r}_m)\}$ are independent.

A.3 Analysis of the signal

Let us assume the signal is bandwidth limited. Let the Fourier transform of the cross term, $X = 2\sqrt{RT}\cos(\mathbf{x}_t \cdot \mathbf{k}_0 + \phi)$, not overlap with the zero-order $R(\mathbf{x}_t) + T(\mathbf{x}_t)$, in k-space. If we write the Fourier transform of the field of the signal as $S(\mathbf{k})$, we will have

$$S(\mathbf{k}) = \sum_{t=1}^{N_P} \left(\sqrt{T(\mathbf{x}_t)}e^{i\phi}\right) e^{-i\mathbf{k}\cdot\mathbf{x}_t}. \tag{A.11}$$

The Fourier transform of the cross term is

$$X(\mathbf{k}) = \sum_{t=1}^{N_P} 2\sqrt{RT(\mathbf{x}_t)}\cos(\mathbf{x}_t \cdot \mathbf{k}_0 - \phi)e^{-i\mathbf{k}\cdot\mathbf{x}_t} = \sqrt{R}S(\mathbf{k}-\mathbf{k}_0) + \sqrt{R}S(\mathbf{k}+\mathbf{k}_0). \tag{A.12}$$

Assume $S(\mathbf{k} - \mathbf{k}_0)$ lies inside the cropped region C, then once we cropped out C and do inverse Fourier transform, the signal is given by

$$\hat{S}(\mathbf{r}_m) = \frac{1}{N_C}\sum_{j=1}^{N_C} \sqrt{R}S(\mathbf{k}_j - \mathbf{k}_0)e^{i\mathbf{r}_m\cdot\mathbf{k}'_j} = \frac{1}{N_C}\sum_{j=1}^{N_C} \sqrt{R}S(\mathbf{k}_j)e^{i\mathbf{r}_m\cdot\mathbf{k}_j}e^{i(\mathbf{k}_0-\mathbf{k}'_0)\cdot\mathbf{r}_m}. \tag{A.13}$$

A.4 SNR calculation

The final signal we will get using this cropping method is

$$\text{Signal}(\mathbf{r}_m) = \hat{S}(\mathbf{r}_m) + \hat{N}(\mathbf{r}_m) = \hat{S}(\mathbf{r}_m) + A(\mathbf{r}_m)e^{i\alpha(\mathbf{r}_m)}. \tag{A.14}$$

So the squared norm of the signal is:

$$|\text{Signal}(\mathbf{r}_m)|^2 = |\hat{S}(\mathbf{r}_m)|^2 + 2|\hat{S}(\mathbf{r}_m)| \cdot A(\mathbf{r}_m)\cos(\alpha'(\mathbf{r}_m)) + A^2(\mathbf{r}_m), \tag{A.15}$$

where $A^2(\mathbf{r}_m) \sim \text{Exponential}(\frac{N_C}{2\sigma'^2})$, and $\alpha'(\mathbf{r}_m) = \alpha(\mathbf{r}_m) - Arg\left[\hat{S}(\mathbf{r}_m)\right]$ which is uniformly distributed $\alpha'(\mathbf{r}_m) \sim \text{Unif}[0, 2\pi]$. Note that the argument of $\hat{S}(\mathbf{r}_m)$ is deterministic (i.e., the signal is deterministic in one experiment, the only variant is noise) so α' is uniformly distributed.

Let us consider $A(\mathbf{r}_m)\cos(\alpha'(\mathbf{r}_m))$, the expected value is

$$\mathbb{E}\left[A(\mathbf{r}_m)\cos(\alpha'(\mathbf{r}_m))\right] = \mathbb{E}\left[A(\mathbf{r}_m)\right]\mathbb{E}\left[\cos(\alpha'(\mathbf{r}_m))\right] = 0. \tag{A.16}$$

Its variance is given by

$$\mathbb{E}\left[A^2(\mathbf{r}_m)\cos^2(\alpha'(\mathbf{r}_m))\right] = \mathbb{E}\left[A^2(\mathbf{r}_m)\right]\mathbb{E}\left[\cos^2(\alpha'(\mathbf{r}_m))\right] = \frac{\sigma^2 N_P}{2N_C}. \tag{A.17}$$

So the mean of $|\text{Signal}|^2$ is

$$\mathbb{E}[|\text{Signal}(\mathbf{r}_m)|^2] = |\hat{S}(\mathbf{r}_m)|^2 + \frac{2\sigma'^2}{N_C} = |\hat{S}(\mathbf{r}_m)|^2 + \frac{\sigma^2 N_P}{N_C}. \tag{A.18}$$

The variance of $|\text{Signal}|^2$ is

$$\begin{aligned}\text{Var}(|\text{Signal}|^2) &= \mathbb{E}\left[\left|2\hat{S}(\mathbf{r}_m)A(\mathbf{r}_m)\cos(\alpha'(\mathbf{r}_m)) + A^2(\mathbf{r}_m)\right|^2\right] \\ &\qquad - \mathbb{E}^2\left[2|\hat{S}(\mathbf{r}_m)|A(\mathbf{r}_m)\cos(\alpha'(\mathbf{r}_m)) + A^2(\mathbf{r}_m)\right] \\ &= 4|\hat{S}(\mathbf{r}_m)|^2\frac{\sigma^2 N_P}{2N_C} + \frac{2\sigma^4 N_P^2}{N_C^2} - \frac{\sigma^4 N_P^2}{N_C^2} \\ &= 2|\hat{S}(\mathbf{r}_m)|^2\frac{\sigma^2 N_P}{N_C} + \frac{\sigma^4 N_P^2}{N_C^2}.\end{aligned} \tag{A.19}$$

So, the SNR is given by

$$\text{SNR}(\mathbf{r}_m) = \frac{|\hat{S}(\mathbf{r}_m)|^2}{\sqrt{2|\hat{S}(\mathbf{r}_m)|^2\frac{\sigma^2 N_P}{N_C} + \frac{\sigma^4 N_P^2}{N_C^2}}}. \tag{A.20}$$

To verify the theoretical SNR, a simulation was conducted. Poisson noise was added to the same speckle field and reference. And this process was repeated for 500 times to calculate the standard deviation of the noise on each pixels. The correlation factor between theory and simulation is 0.998.

The overall SNR (i.e., SNR for the power on the entire detector) is

$$\text{SNR}_{all} = \frac{\sum_{m=1}^{N_C}|\hat{S}(\mathbf{r}_m)|^2}{\sqrt{\sum_{m=1}^{N_C}2|\hat{S}(\mathbf{r}_m)|^2\frac{\sigma^2 N_P}{N_C} + N_C\frac{\sigma^4 N_P^2}{N_C^2}}} = \frac{\sqrt{N_C}\mu_S}{\sqrt{2\mu_S\frac{\sigma^2 N_P}{N_C} + \frac{\sigma^4 N_P^2}{N_C^2}}}, \tag{A.21}$$

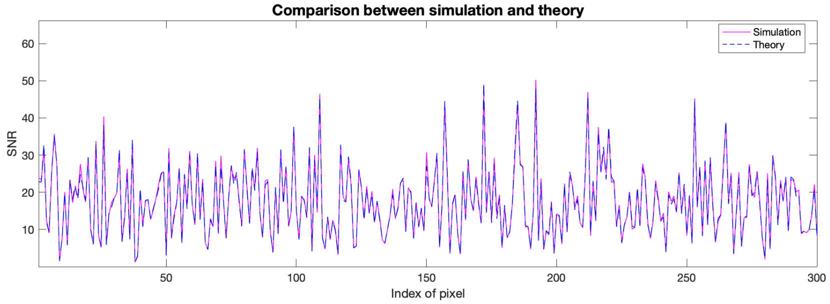

Figure A.3: SNR in simulation vs theoretical prediction in offaxis. In simulation, Poisson noise is added.

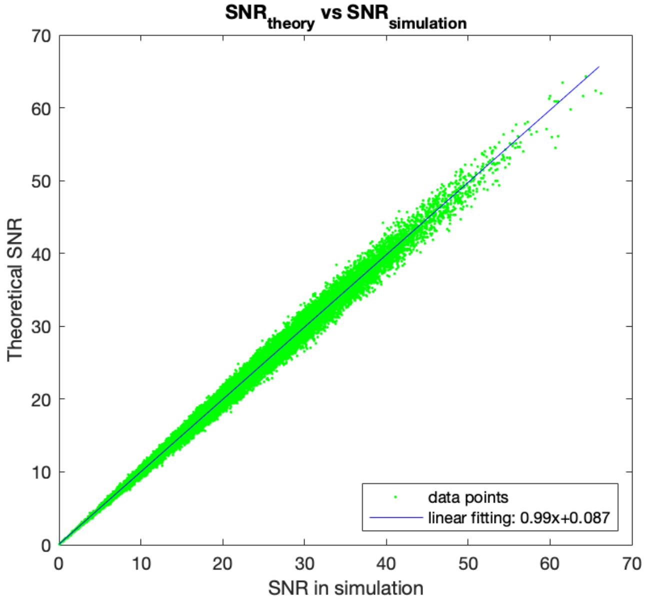

Figure A.4: Linear fitting of offaxis SNR in theory vs in simulation.

where $\mu_S = \frac{1}{N_C}\sum_{m=1}^{N_C} |\hat{S}(\mathbf{r}_m)|^2$. (Note that after cropping, there is only N_C pixels).

Figure A.5 is the comparison between simulation and theory. The noises are simulated for 500 times to calculate their standard deviations.

A.5 The non-interfering part

Let us consider the case that there is a field that does not interfere with the reference and the signal arm. One of such example can be seen in ultrasound optical tomography (UOT), where the untagged light is frequency shifted and the cross-term will be averaged out during the integration time of the detector. Let the non-interfering part at $\mathbf{x}_t$ be a non-negative random variable $X(\mathbf{x}_t)$. Let us plug this into Eq. A.1,

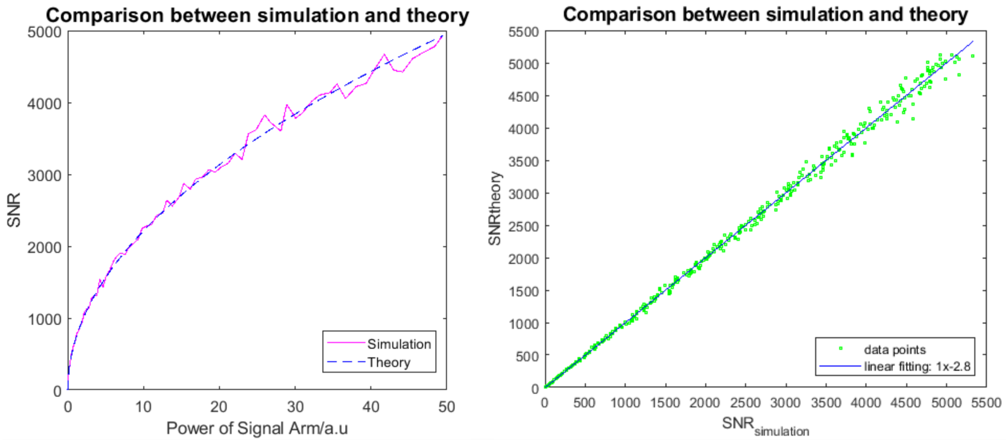

Figure A.5: Overall offaxis SNR. Here, the reconstructed intensities in the entire sensor are added together. This gives us the total power of the ultrasound tagged light. The overall SNR denotes the SNR of this total power.

which gives

$$\text{Offaxis}(\mathbf{x}_t) = R(\mathbf{x}_t) + T(\mathbf{x}_t) + X(\mathbf{x}_t) + 2\sqrt{RT(\mathbf{x}_t)}\cos(\mathbf{k}_0 \cdot \mathbf{x}_t - \phi) + N(\mathbf{x}_t). \quad \text{(A.22)}$$

As $R(\mathbf{x}_t)$ is large and $X(\mathbf{x}_t)$ is non-negative, the noise can be also modeled as a Gaussian random variable with mean zero and variance $R(\mathbf{x}_t) + X(\mathbf{x}_t)$. For one realizations, $X(\mathbf{x}_t)$ is deterministic and we will treat $X(\mathbf{x}_t)$ as deterministic. Given $X(\mathbf{x}_t)$, the noise can be written as

$$N'(\mathbf{x}_t) := N(\mathbf{x}_t)|X(\mathbf{x}_t) \sim \mathcal{N}\left(0, X(\mathbf{x}_t) + R\right). \quad \text{(A.23)}$$

The Fourier transform of the noise term can be written as

$$N'(\mathbf{k}) = \sum_{t=1}^{N_P} N'(\mathbf{x}_t) e^{-i\mathbf{k}\cdot\mathbf{x}_t} = \sum_{t=1}^{N_P} N'(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t) - i\sum_{t=1}^{N_P} N'(\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t). \quad \text{(A.24)}$$

The covariance between the imaginary and real part of $N'(\mathbf{k})$ is

$$\begin{aligned}
\mathbb{E}\Big[-\Big(\sum_{t=1}^{N_P} N'(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\Big)\Big(\sum_{t}^{N_P} N'(\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t)\Big)\Big]& \\
&= -\mathbb{E}\Big[\sum_{t=1}^{N_P} N'^2(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t)\Big] \\
&= -\sum_{t=1}^{N_P}\mathbb{E}\left[N'^2(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t)\right] \\
&= -\sum_{t=1}^{N_P}[R+X(\mathbf{x}_t)]\cos(\mathbf{k}\cdot\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t) \\
&= -\sum_{t=1}^{N_P}[R+X(\mathbf{x}_t)]\frac{1}{2}\sin(2\mathbf{k}\cdot\mathbf{x}_t) \\
&= \Im\Big[\frac{1}{2}\sum_{t=1}^{N_P}X(\mathbf{x}_t)e^{-i(2\mathbf{k})\cdot\mathbf{x}_t}\Big] \\
&= \frac{1}{2}\Im\left[\tilde{X}(2\mathbf{k})\right], \qquad \text{(A.25)}
\end{aligned}$$

where $\tilde{X}$ is the Fourier transform of X. If X is band-limited, which is the usual case in off-axis, we will write k_c as the cut-off frequency of X. Let $\mathbf{k}_m$ be the maximum frequency in DFT, then the covariance of the imaginary and real part of $N'(\mathbf{k})$ is non-zero when $\mathbf{k}\in Z := \{\mathbf{k} \mid \|2\mathbf{k}-2n\mathbf{k}_m\|_2 \le k_c, n\in\mathbb{Z}\}$.

Let us calculate the variance of the real part of $N'(\mathbf{k})$:

$$\begin{aligned}
\mathrm{Var}\Big[\sum_{t=1}^{N_P}N'(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\Big] &= \mathbb{E}\Bigg[\Bigg(\sum_{t=1}^{N_P}N'(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\Bigg)^2\Bigg] \\
&= \sum_{t=1}^{N_P}\mathbb{E}\left[N'^2(\mathbf{x}_t)\cos^2(\mathbf{k}\cdot\mathbf{x}_t)\right] \\
&= \sum_{t=1}^{N_P}\mathbb{E}\Big[\frac{1}{2}N'^2(\mathbf{x}_t)[\cos(2\mathbf{k}\cdot\mathbf{x}_t)+1]\Big] \\
&= \sum_{t=1}^{N_P}\frac{1}{2}\left[R(\mathbf{x}_t)+X(\mathbf{x}_t)\right]\left[\cos(2\mathbf{k}\cdot\mathbf{x}_t)+1\right] \\
&= \frac{1}{2}\Re\left[\tilde{X}(2\mathbf{k})\right]+\frac{1}{2}\left[\tilde{R}(\mathbf{0})+\tilde{X}(\mathbf{0})\right]. \qquad \text{(A.26)}
\end{aligned}$$

Similarly, the variance of the imaginary part of $N'(\mathbf{k})$ is

$$\begin{aligned}
&\mathrm{Var}\left[\sum_{t=1}^{N_P} N'(\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t)\right] = \mathbb{E}\left[\left(\sum_{t=1}^{N_P} N'(\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t)\right)^2\right] \\
&= \sum_{t=1}^{N_P}\mathbb{E}\left[N'^2(\mathbf{x}_t)\sin^2(\mathbf{k}\cdot\mathbf{x}_t)\right] = \sum_{t=1}^{N_P}\mathbb{E}\left[\frac{1}{2}N'^2(\mathbf{x}_t)\left(-\cos(2\mathbf{k}\cdot\mathbf{x}_t)+1\right)\right] \\
&= \sum_{t=1}^{N_P}\frac{1}{2}\left[R(\mathbf{x}_t)+X(\mathbf{x}_t)\right]\left[-\cos(2\mathbf{k}\cdot\mathbf{x}_t)+1\right] \\
&= -\frac{1}{2}\Re\left[\tilde{X}(2\mathbf{k})\right]+\frac{1}{2}\left[\tilde{R}(\mathbf{0})+\tilde{X}(\mathbf{0})\right]. \qquad \text{(A.27)}
\end{aligned}$$

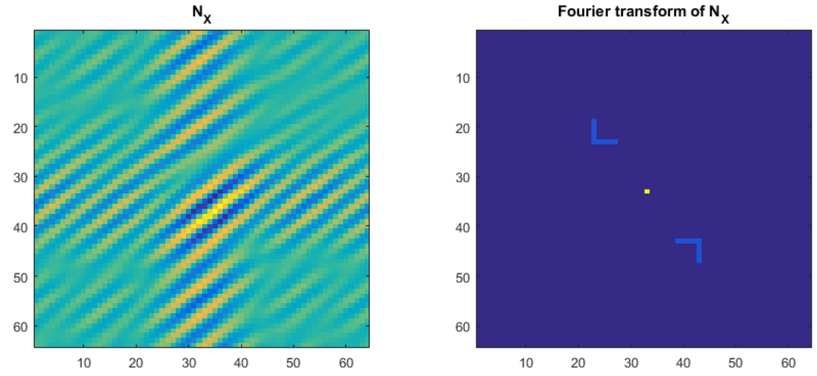

Figure A.6: The image of the non-interfering part and its Fourier transform. In this simulation, the ‘L’ shape in the Fourier transform is chosen deliberately to better visualize the correlation in the following section. The Fourier transform on the right shows only the amplitude.

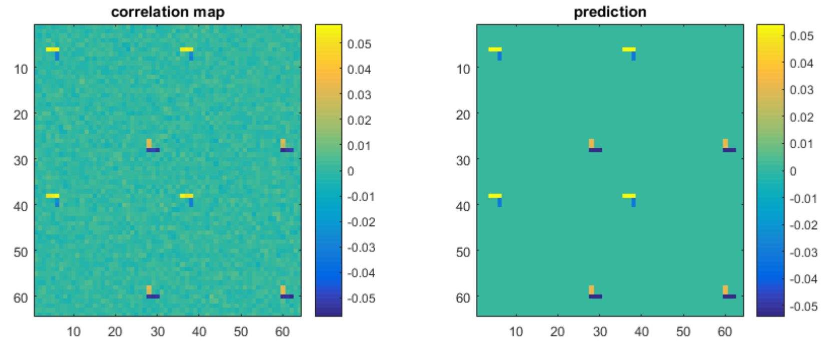

Figure A.7: Correlation of the real and imaginary part of the noise. Left: simulation, right: prediction.

To better quantify the dependence between the imaginary and real part, we will use the correlation between them ($\mathbf{k} \neq \mathbf{0}$)

$$\begin{aligned}\operatorname{Corr}\left[\Re\left(N'(\mathbf{k})\right), \Im\left(N'(\mathbf{k})\right)\right] &= \frac{\operatorname{Cov}\left[\Re\left(N'(\mathbf{k})\right), \Im\left(N'(\mathbf{k})\right)\right]}{\sqrt{\operatorname{Var}\left[\Re\left(N'(\mathbf{k})\right)\right]\operatorname{Var}\left[\Im\left(N'(\mathbf{k})\right)\right]}} \\ &= \frac{\Im\left[\tilde{X}(2\mathbf{k})\right]}{\sqrt{\left[\tilde{R}(\mathbf{0}) + \tilde{X}(\mathbf{0})\right]^2 - \Re\left[\tilde{X}(2\mathbf{k})\right]^2}}. \end{aligned} \tag{A.28}$$

Figures A.6 and A.7 are the simulation results of adding a artificial X term. 4000 samples were simulated in each trial and then the correlation was calculated using the 4000 samples generated.

For a typical speckle field, the imaginary and real part of $\tilde{X}(\mathbf{k})$ ($\mathbf{k} \neq \mathbf{0}$) is usually much smaller than $\tilde{X}(\mathbf{0})$, and the ratio also decreases when the number of $|\mathbf{k}| < k_c$ increases. So for many X (e.g., speckles), the real and imaginary part of $N'(\mathbf{k})$ may be treated as two independent random variables.

Apart from the dependence of the real and imaginary part, we can also find that $N'(\mathbf{k})$ and $N'(\mathbf{k}')$ are dependent even if $\mathbf{k} \neq \mathbf{k}'$.

The covariance between the real part of $N'(\mathbf{k})$ and $N'(\mathbf{k}')$ can be written as

$$\begin{aligned}&\mathbb{E}\left[\left(\sum_{t=1}^{N_P} N'(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\right)\left(\sum_{t}^{N_P} N'(\mathbf{x}_t)\cos(\mathbf{k}'\cdot\mathbf{x}_t)\right)\right] \\ &\qquad = \mathbb{E}\left[\sum_{t=1}^{N_P} N'^2(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\cos(\mathbf{k}'\cdot\mathbf{x}_t)\right] \\ &\qquad = \sum_{t=1}^{N_P} \mathbb{E}\left[N'^2(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\cos(\mathbf{k}'\cdot\mathbf{x}_t)\right] \\ &\qquad = \sum_{t=1}^{N_P} \left[R + X(\mathbf{x}_t)\right]\cos(\mathbf{k}\cdot\mathbf{x}_t)\cos(\mathbf{k}'\cdot\mathbf{x}_t) \\ &\qquad = \sum_{t=1}^{N_P} \left[R + X(\mathbf{x}_t)\right]\frac{1}{2}\left(\cos\left[(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}_t\right] + \cos\left[(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}_t\right]\right) \\ &\qquad = \frac{1}{2}\Re\left[\tilde{X}(\mathbf{k}+\mathbf{k}')\right] + \frac{1}{2}\Re\left[\tilde{X}(\mathbf{k}-\mathbf{k}')\right]. \end{aligned} \tag{A.29}$$

The covariance between the real part of $N'(\mathbf{k})$ and the imaginary part of $N'(\mathbf{k}')$ is

$$\begin{aligned}
&\mathbb{E}\left[-\left(\sum_{t=1}^{N_P} N'(\mathbf{x}_t)\cos(\mathbf{k}\cdot\mathbf{x}_t)\right)\left(\sum_{t}^{N_P} N'(\mathbf{x}_t)\sin(\mathbf{k}'\cdot\mathbf{x}_t)\right)\right] \\
&= -\sum_{t=1}^{N_P}[R+X(\mathbf{x}_t)]\frac{1}{2}\left(\sin\left[(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}_t\right]-\sin\left[(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}_t\right]\right) \\
&= \frac{1}{2}\Im\left[\tilde{X}(\mathbf{k}+\mathbf{k}')\right]-\frac{1}{2}\Im\left[\tilde{X}(\mathbf{k}-\mathbf{k}')\right]. \qquad \text{(A.30)}
\end{aligned}$$

The covariance between the imaginary part of $N'(\mathbf{k})$ and$N'(\mathbf{k}')$ is

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{t=1}^{N_P} N'(\mathbf{x}_t)\sin(\mathbf{k}\cdot\mathbf{x}_t)\right)\left(\sum_{t}^{N_P} N'(\mathbf{x}_t)\sin(\mathbf{k}'\cdot\mathbf{x}_t)\right)\right] \\
&= \sum_{t=1}^{N_P}[R+X(\mathbf{x}_t)]\frac{1}{2}\left(\cos\left[(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}_t\right]-\cos\left[(\mathbf{k}+\mathbf{k}')\cdot\mathbf{x}_t\right]\right) \\
&= \frac{1}{2}\Re\left[\tilde{X}(\mathbf{k}-\mathbf{k}')\right]-\frac{1}{2}\Re\left[\tilde{X}(\mathbf{k}+\mathbf{k}')\right]. \qquad \text{(A.31)}
\end{aligned}$$

So actually they are not independent. If we calculate the correlation between them, then the value in the above three equations will be divided by the square root of the product of their variances, which is related to $\tilde{R}(\mathbf{0})+\tilde{X}(\mathbf{0})$.

Again, when the imaginary and real part of $\tilde{X}(\mathbf{k})$ $(\mathbf{k}\neq\mathbf{0})$ is much smaller than $\tilde{X}(\mathbf{0})$, $N'(\mathbf{k})$ and $N'(\mathbf{k}')$ can be regarded as two independent random variables. And if this condition is satisfied, we can just replace σ^2 in Sections A.1-A.4 with $\sigma_2^2 = R+\mu_X$ to calculate the SNR, where μ_X is the mean of X $(\tilde{X}(\mathbf{0}) = \sum_{t=1}^{N_P} X(\mathbf{x}_t) = N_P\mu_X)$. The following two figures show the simulation result and the theoretical prediction which ignores such correlation.

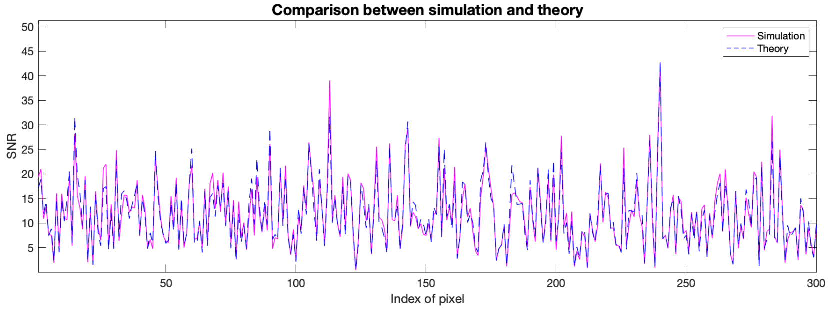

Figure A.8: Pixelwise SNR in theory and simulation (with non-interfering part).

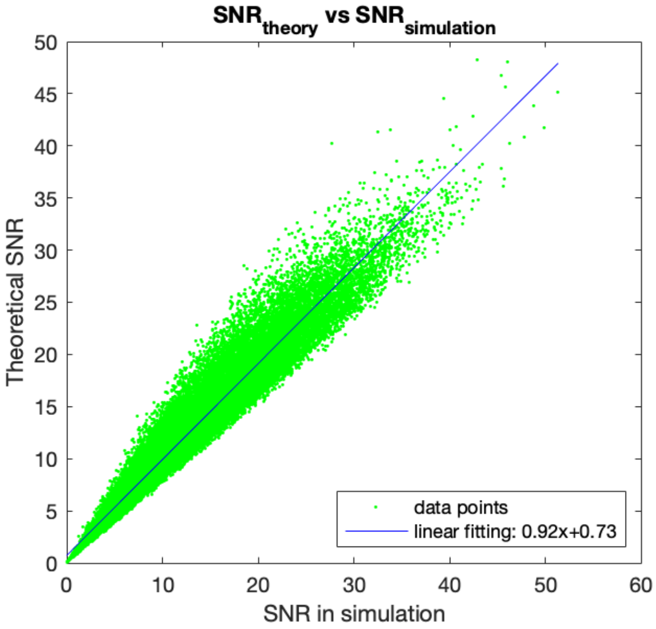

Figure A.9: Linear fitting of simulated pixelwise SNR and prediction (with non-interfering part).

By summing over all pixels, we can also get the power on the detector. Here we did a simulation where the signal power was changed for 150 times. The non-interfering part used in this simulation was speckle patterns. For a fixed signal power, Poisson noise was introduced. Again, this process was repeated for 500 times to obtain the standard deviation of the noise after summation. The theoretical prediction and simulation match pretty well even though the correlation of the noise term is ignored.

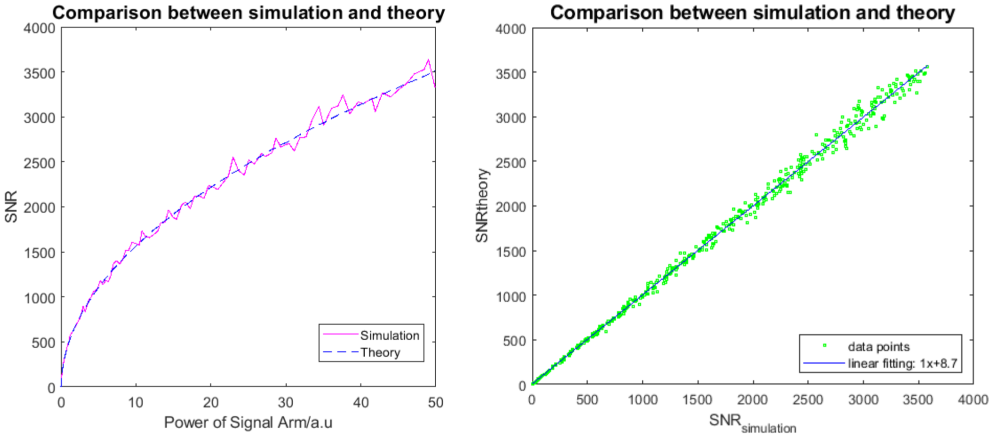

Figure A.10: Overall SNR after summation (with non-interfering part).

A.6 Connecting to SNR in UOT

In UOT, we assumed the mean of X (μ_X) equals R. So, we have $\sigma_2^2 = R + \mu_X = 2R$.

Note the total energy in the cropped out region of the Fourier transform should be the same before and after cropping, we can connect it to the total power after processing ($N_C\mu_S$) via Parseval's theorem. And we can also establish a similar connection to the original off-axis image we get. This gives us

$$N_C(N_C\mu_S) = \text{Power in sideband of FT} = N_P(N_P\bar{N}_T R), \tag{A.32}$$

where $\bar{N}_T = \frac{1}{N_P}\sum_{t=1}^{N_P} T(\mathbf{x}_t)$ is the mean power of the signal arm.

If we use a rectangular aperture to maximize the use of the Fourier space, we will have

$$\frac{N_P}{N_C} = 4. \tag{A.33}$$

Thus, we can conclude

$$\mu_S = \frac{N_P^2}{N_C^2}\bar{N}_T R = 16\bar{N}_T R. \tag{A.34}$$

Substitute what we have now in Eq. A.21, the final SNR expression is

$$\begin{aligned}
\text{SNR}_{\text{UOT}} &= \frac{\sqrt{N_C}\mu_S}{\sqrt{2\mu_S\frac{\sigma_2^2 N_P}{N_C} + \frac{\sigma_2^4 N_P^2}{N_C^2}}} \\
&= \frac{\sqrt{\frac{N_P}{4}}16\bar{N}_T R}{\sqrt{2\cdot 16\bar{N}_T R\cdot 2R\cdot 4 + (2R)^2\cdot 16}} \\
&= \frac{8\sqrt{N_P}\bar{N}_T R}{\sqrt{256\bar{N}_T R^2 + 64R^2}} \\
&= \frac{\sqrt{N_P}\bar{N}_T}{\sqrt{4\bar{N}_T + 1}}.
\end{aligned} \tag{A.35}$$

INDEX