

Computational fluorescence and phase super-resolution microscopy

Li-Hao Yeh



Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2020-36

<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2020/EECS-2020-36.html>

May 1, 2020

Copyright © 2020, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

Computational fluorescence and phase super-resolution microscopy

by

Li-Hao Yeh

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Laura Waller, Chair

Associate Professor Michael Lustig

Associate Professor Na Ji

Spring 2019

The dissertation of Li-Hao Yeh, titled Computational fluorescence and phase super-resolution microscopy, is approved:

Chair	_____	Date	_____
	_____	Date	_____
	_____	Date	_____

University of California, Berkeley

Computational fluorescence and phase super-resolution microscopy

Copyright 2019
by
Li-Hao Yeh

Abstract

Computational fluorescence and phase super-resolution microscopy

by

Li-Hao Yeh

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Associate Professor Laura Waller, Chair

Light microscopy is an important driving force for new biological discoveries because of its capability of visualizing micro-scale cell structures and interactions. Fundamentally, optical resolution is bounded by the diffraction limit, preventing observation of biological events with even smaller scale. However, computational imaging approaches are efficient tools to surpass this limit. In recent years, optimization formulation has become more popular because of its flexibility and efficacy toward information retrieval. In this thesis, we leverage the power of optimization algorithms and new experimental schemes to tackle super-resolution microscopy, both improving existing methods and developing new techniques. We first apply rigorous optimization algorithm analysis to a super-resolution phase microscopy technique, Fourier ptychography. A more accurate noise model and self-calibration algorithm ensure a better reconstruction quality for this technique. Next, we incorporate optimization into the study of a super-resolution fluorescence microscopy technique, structured illumination microscopy. Super-resolution reconstruction is achieved even with a series of random unknown illumination patterns, which is not possible without proper optimization formulation. Next, we leverage the experience of the previous two projects to propose a super-resolution microscopy method for phase and fluorescence contrast with multi-fold resolution improvement in both 2D and 3D using an unknown speckle illumination composed of high-angle plane waves from Scotch tape as a patterning element. The result is a practical method for achieving multimodal super-resolution with $> 2\times$ resolution gain, surpassing the limit of the traditional linear structured illumination microscopy. All these outcomes are within the realm of computational super-resolution microscopy, where the optimization algorithm is jointly designed with optics for efficient information retrieval to achieve super-resolution microscopy.

To mom, who bravely lives her life
To dad, who takes care of the family and leaves me with no worry
To Wen-Ting, who gives me unconditional love and support all along
To science, for being always that fascinating and amazing

Contents

Contents	ii
List of Figures	iv
List of Tables	x
1 Introduction	1
1.1 Diffraction-limited resolution	1
1.2 Super-resolution coherent imaging	6
1.3 Super-resolution incoherent imaging	8
1.4 Computational super-resolution imaging	12
1.5 Outline of thesis	13
2 Optimization analysis for Fourier ptychographic phase microscopy	15
2.1 Introduction	16
2.2 Optimization formulation to Fourier ptychographic microscopy	23
2.3 Algorithms for Fourier ptychographic microscopy	27
2.4 Performance analysis of various algorithms	37
2.5 Joint estimation of pupil function and LED positions	48
2.6 Conclusion	53
3 Computational random structured illumination fluorescence microscopy	55
3.1 Introduction	56
3.2 Theory and method	57
3.3 Simulation results	64
3.4 Experimental results	67
3.5 Reduced acquisitions by multi-spot scanning	68
3.6 Pixel reassignment for better SNR	71
3.7 Extension to 3D imaging	73
3.8 Conclusion	79
4 Computational speckle structured illumination phase and fluorescence microscopy	81

4.1	Introduction	82
4.2	Theory	84
4.3	Experimental results	96
4.4	Discussion	103
4.5	Conclusion	104
5	Computational speckle structured illumination 3D phase and fluorescence microscopy	105
5.1	Introduction	105
5.2	Theory	107
5.3	Experimental results	117
5.4	Discussion	123
5.5	Conclusion	124
6	Conclusion and future work	125
	Bibliography	128

List of Figures

1.1	Diagram of a standard microscope modeled by a $4f$ system, imaging a sample with a transmittance function $o(\mathbf{r})$ illuminated by a coherent wave field $u(\mathbf{r}, t)$	2
1.2	The 2D coherent transfer function and its corresponding point spread function (PSF) with the profile plot showing the function value at the central cutline.	4
1.3	The 2D incoherent transfer function and its corresponding point spread function (PSF) with the profile plot showing the function value at the central cutline.	6
1.4	Comparison of scattered field at the aperture for on-axis illumination and oblique illumination. Coherent super-resolution information (higher-angle scattered light) is brought into the system by oblique illumination.	7
1.5	Comparison of holographic synthetic aperture microscopy and Fourier ptychographic microscopy in terms of Fourier space sampling, type of acquisition, and inversion solver.	8
1.6	Fourier space coverage comparison for a diffraction-limited system, linear SIM under sinusoidal illumination, linear SIM under random illumination, and nonlinear SIM under sinusoidal illumination. The modulation between the structured pattern and the fluorescent sample in real space could be interpreted as a convolution of these variables' spectrum in the Fourier space. The nonlinear SIM is able to achieve larger Fourier space coverage because the nonlinearity helps to create pattern that is much larger than the illumination NA. The inscribed figures in the illumination pattern columns are the corresponding real space images of these illumination patterns.	10
2.1	(a) Experimental setup for Fourier ptychography with an LED array microscope. (b) The sample's Fourier space is synthetically enlarged by capturing multiple images from different illumination angles. Each circle represents the spatial frequency coverage of the image captured by single-LED illumination. Brightfield images have orders of magnitude higher intensity than darkfield (see histograms), resulting in different noise levels.	16
2.2	The full field of view and the zoomed region of the USAF resolution target for the brightfield raw image from central LED illumination and FPM reconstructed image in the central region of the FOV, demonstrating the super-resolution capability.	17

2.3	The full field of view and the zoomed region for both brightfield and FPM reconstructed phase of the Hela cell, respectively.	18
2.4	Schematic of the phase retrieval procedure for the Gerchberg-Saxton (gradient descent) iterative algorithm. The sample's complex field is iteratively updated as its estimate is propagated back and forth between real space and Fourier space constraints, for each of the measured intensity images taken at multiple Fourier regions.	20
2.5	The general flow of iterative nonlinear optimization algorithms for Fourier ptychographic microscopy.	21
2.6	To explain the artifacts in our experimental results, as well as evaluate the robustness of various algorithms under common types of errors, we simulate several FPM datasets with different types of known error: (1) Ideal data, (2) Poisson noise data, (3) aberrated data, (4) LED misaligned data (\times : original position, \circ : perturbed position).	38
2.7	Reconstructed amplitude from simulated datasets with three types of errors, using different algorithms. The intensity-based algorithms suffer from high frequency artifacts under both noise and model mis-match errors. The percentage on the top left corner of each image is the relative error of each reconstruction.	39
2.8	Reconstructed phase from simulated datasets with three types of errors, using different algorithms. The intensity-based algorithms suffer from phase wrapping artifacts under both noise and model mis-match errors. The percentage on the top left corner of each image is the relative error of each reconstruction.	40
2.9	Fourier ptychographic reconstruction (amplitude only) of a test object with the algorithms discussed here, all using the same experimental dataset. Algorithms derived from the same cost function (amplitude-based, intensity-based, and Poisson-likelihood) give similar performance, and first-order methods (Gerchberg-Saxton) suffer artifacts.	41
2.10	Phase relative error as a function of iteration number for different algorithms with the (a) ideal data, (b) Poisson noise data, (c) aberrated data and (d) LED misaligned data. When the data is not perfect, some of the algorithms may not converge to a correct solution.	43
2.11	Both Poisson noise and model mis-match (aberrations, LED misalignment) cause errors that scale with mean intensity. Here, histograms show the intensity deviations under Poisson noise, aberration, and misalignment for a brightfield and darkfield image.	46
2.12	The <i>intensity-based</i> cost function gives higher weighting to images in the low spatial frequency region of the Fourier domain, resulting in high-frequency artifacts. Here, we show the gradient of the <i>amplitude-based</i> , <i>Poisson-likelihood-based</i> and <i>intensity-based</i> cost functions at the tenth iteration, using experimental data. . .	47
2.13	The flow of model mis-match correction algorithms for Fourier ptychographic microscopy. Redundancy and diversity in the dataset enables algorithmic self-calibration of aberrations (pupil function) and LED position errors.	49

2.14	Object and pupil reconstruction results using different algorithms, with and without pupil estimation. The second-order method (sequential Gauss-Newton) with pupil estimation gives the best result, as expected. In this case, we find that the second-order method <i>without</i> pupil estimation is already better than first-order method (sequential gradient descent) <i>with</i> pupil estimation.	50
2.15	(a) Adding LED misalignment correction improves the reconstruction results (sequential Gauss-Newton method). (b) The original, perturbed, and corrected LED positions in angular coordinates. LED correction accurately retrieves the actual LED positions.	51
2.16	Experimental reconstructions with and without LED misalignment correction (sequential Gauss-Newton method). (a) The reconstructed object and pupil. (b) The original and corrected LED positions, in angular coordinates.	53
3.1	Example experimental setup for structured illumination microscopy (SIM) using a deformable mirror device (DMD) to capture low-resolution images of the object modulated by different illumination patterns. Our IPE-SIMS algorithm reconstructs both the super-resolved image and the unknown arbitrary illumination patterns.	58
3.2	The first part of our algorithm, Pattern Estimation (PE), iteratively estimates the illumination patterns from an approximated object given by the deconvolved widefield image.	59
3.3	The second part of our algorithm, termed structured illumination microscopy with a statistical prior (SIMS), estimates the high-resolution object from the measured images and the estimated illumination patterns obtained in Part 1.	61
3.4	(a) Simulated reconstructions of a Siemens star target under a widefield microscope, deconvolved widefield, confocal microscope, deconvolved confocal, blind SIM [51], S-SOFI [55], our PE-SIMS and PE-SIMS-PR algorithms. (b) The effective modulation transfer function (MTF) of each method, given by the contrast of the reconstructed Siemens star image at different radii.	66
3.5	Reconstructions of red fluorescent beads (Ex:580 nm/Em:605 nm) from the experiment using random pattern illumination with 20×20 scanning step.	68
3.6	Comparison of our algorithm on dataset from Multispot SIM (MSIM) which uses with $N_{\text{img}} = 224$ scanned multi-spot patterns from [45]. We show the deconvolved widefield image and the reconstructions using MSIM with known patterns, as well as our blind PE-SIMS algorithm with and without pixel reassignment.	69
3.7	Results with simulated and experimental (fluorescent beads) datasets comparing random speckle and multi-spot illumination patterns. (middle row) Shading maps overlaid on the object. Decreasing the number of random patterns results in shading artifacts in the reconstruction. The random patterns are scanned in 20×20 , 10×10 , and 6×6 steps with the same step size of 0.6 FWHM of the PSF, while the multi-spot pattern is scanned with 6×6 steps.	70

3.8	(a) Comparison of the PSF and OTF for SIMS and SIMS with pixel reassignment (PR). (b) Comparisons of the deconvolved widefield image and the reconstructions of the 6×6 multi-spot scanned fluorescent beads with and without pixel reassignment.	73
3.9	(Left) x - y and x - z cross-sections of the experimental raw data from fluorescent beads illuminated with a multispot pattern. (Right) The estimated 3D illumination pattern from our pattern estimation algorithm.	77
3.10	(Top) Comparison of widefield and 3D PE-SIMS reconstruction, showing x - y and x - z cross-sections for a sample with 200 nm red fluorescent beads. (Bottom) Histogram on the size (FWHM) of 32 reconstructed fluorescent beads.	78
3.11	Comparison of widefield and 3D PE-SIMS reconstruction, showing x - y and x - z cross-sections for a 3D bacteria aggregation sample (<i>Candidatus Kuenenia stuttgartiensis</i>).	79
4.1	Structured illumination microscopy (SIM) with laterally-translated Scotch tape as the patterning element, achieving $4\times$ resolution gain. Our imaging system has both an incoherent arm, where Sensor-F captures raw fluorescence images (at the emission wavelength, $\lambda_{em} = 605$ nm) for fluorescence super-resolution, and a coherent arm, where Sensor-C1 and Sensor-C2 capture images with different defocus (at the laser illumination wavelength, $\lambda_{ex} = 532$ nm) for both super-resolution phase reconstruction and speckle trajectory calibration. OBJ: objective, AP: adjustable iris-aperture, DM: dichroic mirror, SF: spectral filter, ND-F: neutral-density filter.	84
4.2	Verification of fluorescence super-resolution with $4\times$ resolution gain. Widefield images, for comparison, were acquired at (a) 0.1 NA and (e) 0.4 NA by adjusting the aperture size. (b) The Scotch tape speckle pattern creates much higher spatial frequencies (~ 0.35 NA) than the 0.1 NA detection system can measure. (c) Using the 0.1 NA aperture, we acquire low-resolution fluorescence images for different lateral positions of the Scotch tape. (d) The reconstructed SIM image contains spatial frequencies up to ~ 0.4 NA and is in agreement with (e) the deconvolved widefield image with the system operating at 0.4 NA.	97
4.3	Verification of coherent quantitative phase (QP) super-resolution with $4\times$ resolution gain. (a) Low-resolution intensity image and (b) “ground truth” phase at NA=0.4, for comparison. (c) Raw acquisitions of the speckle-illuminated sample intensity from two focus planes, collected with 0.1 NA. (d) Reconstructed SR amplitude and QP, demonstrating $4\times$ resolution gain.	98
4.4	Reconstructed super-resolution fluorescence with $4\times$ resolution gain across the full FOV (See Visualization 1). Four zoom-ins of regions-of-interest (ROIs) are compared to their widefield counterparts.	99

4.5	Reconstructed multimodal (fluorescence and quantitative phase) high-content imaging (See Visualization 2 and 3). Zoom-ins for three ROIs compare the wide-field, super-resolved fluorescence, coherent intensity, and super-resolved phase reconstructions.	100
4.6	Algorithmic self-calibration significantly improves fluorescence super-resolution reconstructions. Here, we compare the reconstructed fluorescence image, speckle intensity, and OTF with no correction, OTF correction, and both OTF correction and scanning position correction. The right panel shows the overlay of the uncorrected and corrected scanning position trajectories.	102
4.7	Algorithmic self-calibration significantly improves coherent super-resolution reconstructions. We show a comparison of reconstructed amplitude, phase, speckle amplitude, and phase of the pupil function with no correction, pupil correction, and both pupil correction and scanning position correction. The right panel shows the overlay of scanning position trajectory for the in-focus and defocused cameras before and after correction.	103
5.1	3D multimodal structured illumination microscopy (SIM) with laterally translating Scotch tape as the patterning element. The coherent arm (Sensor-C1 and Sensor-C2) simultaneously captures images with different defocus at the laser illumination wavelength ($\lambda_{\text{ex}} = 532 \text{ nm}$), used for both 3D phase retrieval and speckle trajectory calibration. The incoherent (fluorescence) arm (Sensor-F) captures low-resolution raw fluorescence acquisitions at the emission wavelength ($\lambda_{\text{em}} = 605 \text{ nm}$) for 3D fluorescence super-resolution reconstruction. OBJ: objective, DM: dichroic mirror, SF: spectral filter, ND-F: neutral-density filter.	107
5.2	3D coherent and incoherent transfer function (TF) analysis of the SIM imaging process. The 3D (a) coherent and (b) incoherent TFs of the detection system are autocorrelated with the 3D Fourier support of the (c) illumination speckle field and (d) illumination intensity, respectively, resulting in the effective Fourier support of 3D (e) coherent and (f) incoherent SIM. In (e) and (f), we display decomposition of the auto-correlation in two steps: ① tracing the illumination support in one orientation and ② replicating this trace in azimuthal direction.	108
5.3	3D multi-slice model: (a) coherent and (b) incoherent imaging models for interaction between the sample and the speckle field.	109
5.4	3D multimodal (fluorescence and phase) SIM reconstruction compared to wide-field fluorescence and coherent intensity images for 700 nm fluorescent microspheres. Resolution beyond the system's diffraction limit is achieved in both the (a) coherent and (b) fluorescent arms.	119
5.5	Reconstructed 3D multimodal (fluorescence and phase) large-FOV for mixed 2 μm , 4 μm fluorescent and 3 μm non-fluorescent polystyrene microspheres. Zoom-ins for two ROIs show fluorescence and phase at different depths.	121

5.6	Reconstructed 3D multimodal (fluorescence and phase) large-FOV imaging for HT-29 cells. Zoom-ins for two ROIs show fluorescence and phase at different depths. The blue arrows in two ROIs indicate two-layer cell clusters that come in and out of focus. The orange arrows indicate intracellular components, including higher-phase-contrast lipid vesicles at $z = -5.1 \mu\text{m}$, nucleolus at $z = 0$, as well as the cell nucleus and cell-cell membrane contacts.	122
-----	--	-----

List of Tables

2.1	Tuning Parameters	42
2.2	Convergence Speed	44
3.1	Achieved resolution for different algorithms	65
5.1	Summary of spatial frequency bandwidths	109

Acknowledgments

First of all, I would like to express my sincerest gratitude to my advisor, Professor Laura Waller. She is not only a pure academic advisor for research discussion but also a life role model that demonstrates the importance of caring and supporting in academia. Her nurturing guiding style has played an important role in shaping my great PhD experiences. I have learned so much from her in many occasions such as our weekly meetings, the time we went on the same conferences, the time we worked together on papers, and the time we chatted over casual social events. In meetings, I was always inspired by her intuitive optics understanding and clear presentation of ideas. On conferences and social events, she always encourages me to promote myself and our research to more people in the field. Working together on papers, I am amazed by her diligent working attitude and meticulousness in paragraphs and wording to show the best clarity of ideas. There are so much more I can list about my learning from Laura.

An important perspective of graduate school to me is teaching and I would really like to thank Professor Michael Lustig for offering me the opportunity to teach a core EE class EE123 Digital Signal Processing for two semesters. He made me realize how to contribute a class in order to let student have an enjoyable time. Working with him on the class has been a great memory during my time in Berkeley. Besides class, I also benefitted from occasional discussion with him on my research project. He always gives incredible insights from the signal processing prospective. I am really grateful to have him on my thesis committee.

I would really like to thank my other thesis committee, Professor Na Ji, for providing valuable advice over my qualifying exam and some other occasions. I am inspired by her insight in solving difficult microscopy problems.

A great working environment is key to great PhD experience. Thus, I am really grateful to be a member of Computational Imaging Lab in Berkeley. I really enjoy the conversation with passionate and friendly people in the office, including Emrah, Volker, Regina, Emma, Yi, Stuart, and Henry. I also love the company with Kyrollos and Linda in the lab over countless late nights. Occasionally, running into Herbert, Gautam, David, Grace, Michael Kellman, and Kristina over the weekend cheers my tiring mind and keeps me working for a longer run. I cherish the experience working closely with Michael Chen in my first year, where we bounce ideas around and learn phase imaging together. I am grateful that I have chance to work with Nick when teaching EE123, where I learned great leadership and people skills. I would also like to thank my best neighbor Zack for having great conversations about life and research over five years. Especially, I want to thank the inspiring post-docs that I worked with closely over the years. I like to thank Lei for guiding me through the computational imaging field in my first two years. So many conversations in research and in career that really benefit me a lot. I want to thank Jingshan, who teaches me optimization in my first year. I still remember the time that we worked hand in hand on deriving the gradient of FPM with CR calculus. I want to thank Shwetadwip for teaching me great experimental skills and enlightening me in many life and research perspectives. I would

also like to express my gratitude to people that I have worked with throughout the time, including Nicole Repina, Jonathon Dong, Mahdi Soltanolkotabi, Jon Tamir, and Ziji Liu.

Along the PhD journey, I enjoyed the company from my friends in Berkeley. Among them, I would like to first thank Andrew Ho. We came to Berkeley together but with very different career choices. I really appreciate that I have a friend like you all around to share life lessons and grow together. Weekly gathering with my badminton group provides me great amount of support throughout the years as well. Especially, I would like to thank Tsz Him (Zephy) Leung for five years of great conversations over physics and life choices. I am also grateful to have met Siang-Chen Wu to share our academic vision.

In the end, I would like to thank all my family for the unconditional support over these years. Because of their support, I am able to stay focus on my research and reach my achievement today. I would like to specially thank my girlfriend, Wen-Ting (Jocelyn), who recently became my wife. Thanks for her constant support and the encouragement that made me through all the difficulties along the journey.

Chapter 1

Introduction

This thesis shows the progressive building blocks toward computational fluorescence and phase super-resolution microscopy with speckle illumination. This chapter is devoted to the basic concepts required to understand the material in this thesis.

Resolution is one of the important characteristics of a microscope system, specifying the minimum distinguishable feature size. Fundamentally, when the microscope system is free from imperfections, the resolution is limited by the diffraction effect of the light. The so-called diffraction limit was defined 1.5 centuries ago by German Physicist, E. Abbe [1]. This diffraction limit defines the minimum resolvable feature size (*i.e.* the resolution) of a microscope as inversely proportional to the amount of angular information a microscope can collect, and is quantified by numerical aperture (NA). Higher NA leads to better resolution, usually at the cost of small field-of-view (FOV), and limited by physical fabrication constraints and aberration. Breaking the diffraction limit enables observation of features that were previously not resolvable, which could potentially bring new discoveries in biology [2–16]. Lower NA objectives have poor resolution, but usually have larger FOV. Breaking the diffraction limit in this case increases the amount of information transferred through the microscope [17–30], opening up potential applications in high-throughput imaging that is highly desired in the drug industry or system biology [31–37], where scientists need to observe many cells at good resolution.

Throughout this thesis, our discussion is focused on scalar optics, where only one electric field vector component is considered in the calculation.

1.1 Diffraction-limited resolution

In this section, we apply linear system theory to optical imaging systems similar to [38] to analyze the mapping of information. A $4f$ system geometry is chosen for analysis because most microscopes can be simplified into a $4f$ system. Based on the geometry, we will establish the linear input-output relationship for coherent and incoherent imaging scenarios. These input-output relationships are key to understanding diffraction-limited resolution.

1.1.1 Coherent imaging system

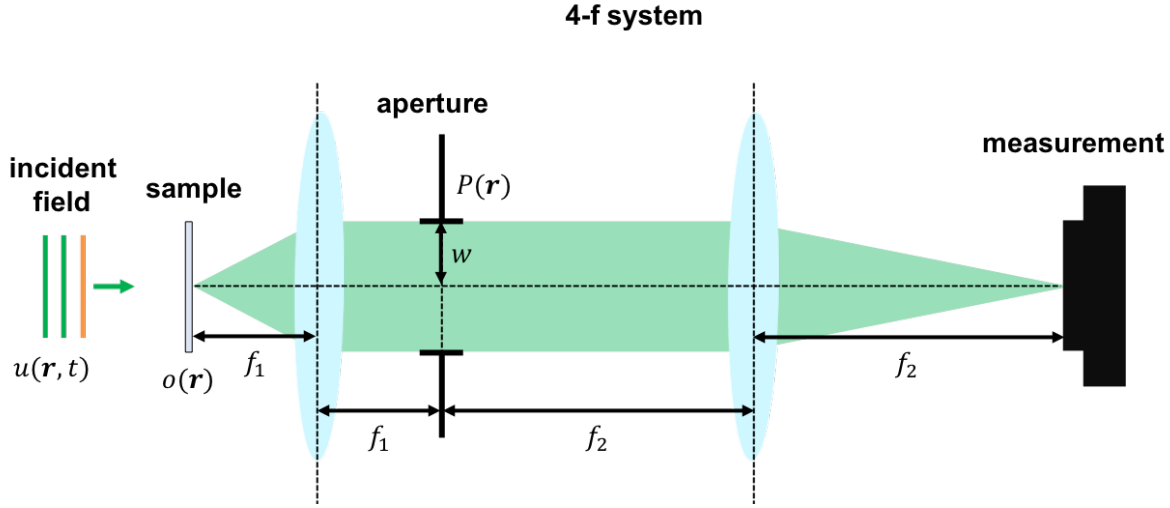


Figure 1.1: Diagram of a standard microscope modeled by a $4f$ system, imaging a sample with a transmittance function $o(\mathbf{r})$ illuminated by a coherent wave field $u(\mathbf{r}, t)$.

Figure 1.1 shows a $4f$ system diagram, representing a typical microscope. With this geometry, we start by establishing the input-output relationship for the coherent imaging case. The coherence we are referring to here is the spatial coherence of the light. To rule out temporal coherence effects, we assume our illumination wave field to be quasi-monochromatic (*i.e.* narrow-band light with temporal coherence length much larger than the spatial coherence length). With this assumption, we represent the incident wave field centered at frequency ν in the time-varying phasor form as

$$U(\mathbf{r}, t) = \mathcal{R} \{ u(\mathbf{r}, t) e^{-i2\pi\nu t} \} = |u(\mathbf{r}, t)| \cos(2\pi\nu t + \angle u(\mathbf{r}, t)), \quad (1.1)$$

where $\mathcal{R}\{\cdot\}$ represents the operation of taking the real part of the variable, $\mathbf{r} = (x, y)$ is the 2D spatial coordinate, and $u(\mathbf{r}, t)$ is the time-varying phasor of the real wave field $U(\mathbf{r}, t)$. In the following discussion, we will use phasor notation to refer to the electric field.

The electric field illuminates a sample whose transmittance function is $o(\mathbf{r})$ and forms the input wave field

$$u_i(\mathbf{r}, t) = u(\mathbf{r}, t) o(\mathbf{r}). \quad (1.2)$$

According to Chapter 5.2 in [38], the wave field at the aperture plane after going through the first $2f$ system (focal length f_1) is a Fourier transform of the input wave field under paraxial approximation and is expressed as

$$u_a(\mathbf{r}, t) = \frac{1}{i\lambda f_1} \int u_i(\mathbf{r}', t) e^{-\frac{i2\pi\mathbf{r}' \cdot \mathbf{r}}{\lambda f_1}} d\mathbf{r}' = \frac{1}{i\lambda f_1} \tilde{u}_i \left(\frac{\mathbf{r}}{\lambda f_1}, t \right), \quad (1.3)$$

where λ is the central wavelength of the wave field and \tilde{a} represents the 2D spatial Fourier transform of the variable a . Since the wave field at the aperture is the Fourier transform of the input wave field, its distribution represents the angular information of the input wave field. This angular information is truncated by the aperture of the microscope, $P(\mathbf{r})$, which may be complex to model aberration. Thus, a limited amount of information is mapped onto the imaging plane through the second $2f$ system (focal length f_2) to form the output wave field, expressed as

$$\begin{aligned} u_o(\mathbf{r}, t) &= \frac{1}{i\lambda f_2} \int u_a(\mathbf{r}', t) P(\mathbf{r}') e^{-\frac{i2\pi\mathbf{r}'\cdot\mathbf{r}}{\lambda f_2}} d\mathbf{r}' \\ &= -\frac{f_2}{f_1} \int \tilde{u}_i\left(-\frac{f_2\mathbf{u}}{f_1}, t\right) P(-\lambda f_2\mathbf{u}) e^{i2\pi\mathbf{u}\cdot\mathbf{r}} d\mathbf{u}, \quad \mathbf{r}' = -\lambda f_2\mathbf{u} \\ &= u_m(\mathbf{r}, t) \otimes h_c(\mathbf{r}), \end{aligned} \quad (1.4)$$

where \mathbf{u} is the 2D spatial frequency coordinate of the output wave field and \otimes denotes 2D convolution operation. The magnified wave field, $u_m(\mathbf{r}, t)$, the point spread function (PSF) of the coherent imaging system, $h_c(\mathbf{r})$, and system magnification, M , are expressed as

$$\begin{aligned} u_m(\mathbf{r}, t) &= \frac{1}{M} u_i\left(\frac{\mathbf{r}}{M}, t\right), \\ h_c(\mathbf{r}) &= \int P(-\lambda f_2\mathbf{u}) e^{i2\pi\mathbf{u}\cdot\mathbf{r}} d\mathbf{u}, \\ M &= -\frac{f_2}{f_1}. \end{aligned} \quad (1.5)$$

Equation (1.4) has established the linear space-invariant relationship between the input wave field and the output wave field. The output wave field from the $4f$ system is simply a convolution between the magnified input wave field and the PSF of the coherent imaging system, which is a Fourier transform of a scaled pupil function. In the second line of Eq. (1.4), we recognize the transfer function of this linear system to be

$$\tilde{h}_c(\mathbf{u}) = P(-\lambda f_2\mathbf{u}), \quad (1.6)$$

which maps the frequency information between the output wave field and the magnified input wave field (*i.e.* scattered field from the sample), and is called the coherent transfer function. Usually, the aperture function of a non-aberrated system takes the form of a circle function as

$$P(\mathbf{r}) = \begin{cases} 1, & \|\mathbf{r}\|_2 \leq w \\ 0, & \text{otherwise} \end{cases} \Rightarrow \tilde{h}_c(\mathbf{u}) = \begin{cases} 1, & \|\mathbf{u}\|_2 \leq \frac{w}{\lambda f_2} = \frac{\text{NA}}{\lambda} \\ 0, & \text{otherwise} \end{cases} \quad (1.7)$$

where w is the radius of the aperture. This implies the coherent transfer function to be a circle function as well with the radius defined by the ratio of w and f_2 , which is the numerical

aperture (NA) of the system. This coherent transfer function defines the passband of wave-field information. Higher spatial frequency from the input wave field encodes information of smaller feature sizes, which require imaging systems with larger NA to transfer. This is the origin of the diffraction limited resolution for coherent imaging systems. Figure 1.2 shows the coherent transfer function and the corresponding PSF. The profile plot of the PSF demonstrates the full-width-half-maximum of the PSF is around λ/NA , which sets the coherent diffraction limit.

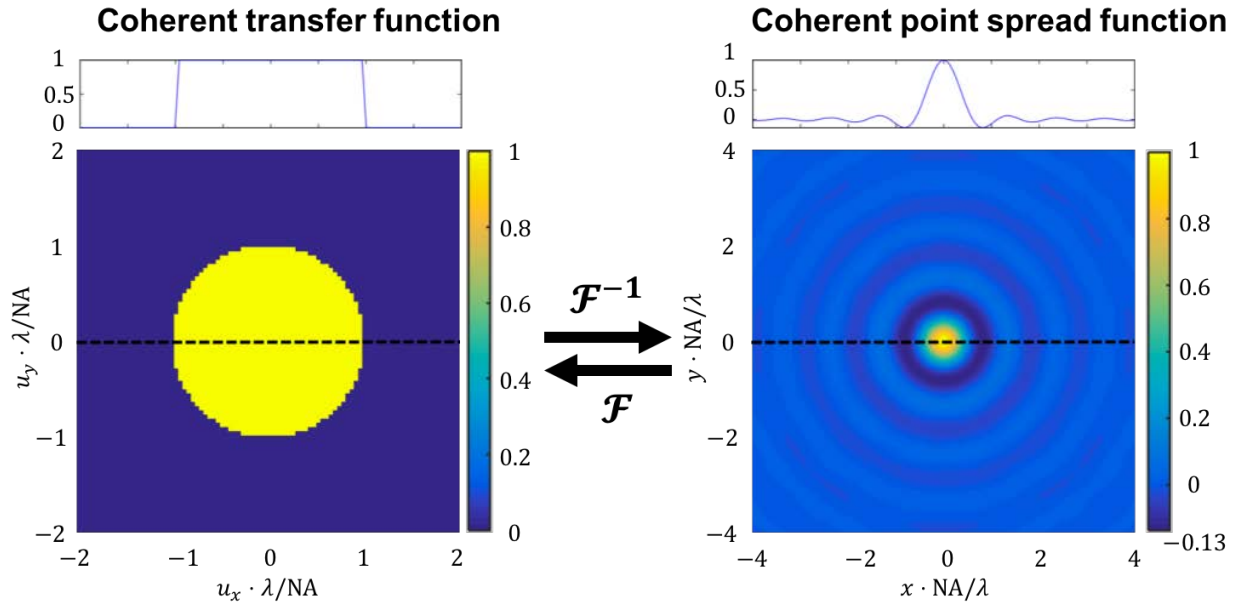


Figure 1.2: The 2D coherent transfer function and its corresponding point spread function (PSF) with the profile plot showing the function value at the central cutline.

Assuming the illumination field is a plane wave from a single coherent light source (e.g. a single laser) and the magnified object is denoted as $o_m(\mathbf{r}) = o\left(\frac{\mathbf{r}}{M}\right)$, we express our measurement at the sensor as the time average of the light intensity

$$\begin{aligned}
 I_o(\mathbf{r}) &= \langle |u_o(\mathbf{r}, t)|^2 \rangle_t = \frac{1}{M^2} \iint \left\langle u\left(\frac{\mathbf{r}_1}{M}, t\right) u^*\left(\frac{\mathbf{r}_2}{M}, t\right) \right\rangle_t o_m(\mathbf{r}_1) o_m^*(\mathbf{r}_2) h_c(\mathbf{r} - \mathbf{r}_1) h_c^*(\mathbf{r} - \mathbf{r}_2) d\mathbf{r}_1 d\mathbf{r}_2 \\
 &= \frac{|u_0|^2}{M^2} \cdot \left| \int o_m(\mathbf{r}_1) h_c(\mathbf{r} - \mathbf{r}_1) d\mathbf{r}_1 \right|^2,
 \end{aligned} \tag{1.8}$$

where we have used

$$\left\langle u\left(\frac{\mathbf{r}_1}{M}, t\right) u^*\left(\frac{\mathbf{r}_2}{M}, t\right) \right\rangle_t = |u_0|^2 \tag{1.9}$$

because the wave field is fully correlated at two points \mathbf{r}_1 and \mathbf{r}_2 [39]. The measured intensity is nonlinear to the input wave field (here the magnified transmittance function). If we are able

to extract the output wave field information using holography or a phase retrieval method, we can use the previously established input-output wave field relationship from Eq. (1.4) for resolution analysis.

1.1.2 Incoherent imaging system

Similar to the coherent imaging scenario, we start the resolution analysis for an incoherent imaging system by building up the input-output relationship for a $4f$ geometry. The fundamental wave propagation physics is the same as what we derived previously in the coherent imaging case. The major difference happens in Eq. (1.8). For an incoherent imaging system, the correlation of the wave field at position \mathbf{r}_1 and \mathbf{r}_2 is 0 except for $\mathbf{r}_1 = \mathbf{r}_2$, where we could express it as

$$\langle u(-\mathbf{r}_1, t)u^*(-\mathbf{r}_2, t) \rangle_t = \kappa\delta(\mathbf{r}_1 - \mathbf{r}_2), \quad (1.10)$$

where κ is a constant in the unit of power to balance out the unit of a delta function such that the overall unit of the field correlation is intensity. This leads us to write out the measurement for an incoherent imaging system as

$$I_o(\mathbf{r}) = \frac{\kappa}{M^2} \int |o_m(\mathbf{r}_1)|^2 |h_c(\mathbf{r} - \mathbf{r}_1)|^2 d\mathbf{r}_1. \quad (1.11)$$

Note that this is a linear space-invariant relationship with respect to intensity. The output intensity is now a convolution of input intensity (absolute square of the magnified transmittance function) and the intensity of the coherent PSF. The input-output relationship in the incoherent imaging system is built on the intensity, not the wave field as in coherent imaging case.

With this understanding, we identify the incoherent PSF as

$$h_{inc}(\mathbf{r}) = |h_c(\mathbf{r})|^2. \quad (1.12)$$

The Fourier transform of the incoherent PSF is thus

$$\tilde{h}_{inc}(\mathbf{u}) = \tilde{h}_c \star \tilde{h}_c(\mathbf{u}), \quad (1.13)$$

where \star is the autocorrelation operation. Since the incoherent PSF is the absolute square of the coherent PSF, the incoherent transfer function is the autocorrelation of the coherent transfer function. Similar to the coherent imaging case, the incoherent transfer function sets the passband of the intensity information, and hence defines the diffraction-limited resolution for an incoherent imaging system. Figure 1.3 shows the incoherent transfer function and the corresponding PSF. The profile plot of the PSF demonstrates the full-width-half-maximum of the PSF is around $\lambda/2NA$, which is $2\times$ smaller than the coherent diffraction limit. Hence, the intensity resolution with an incoherent system has $2\times$ better resolution than a coherent system. This incoherent imaging model is appropriate for modeling a transmission microscope with an incoherent light source, or a fluorescence microscope, since fluorescent light is also incoherent.

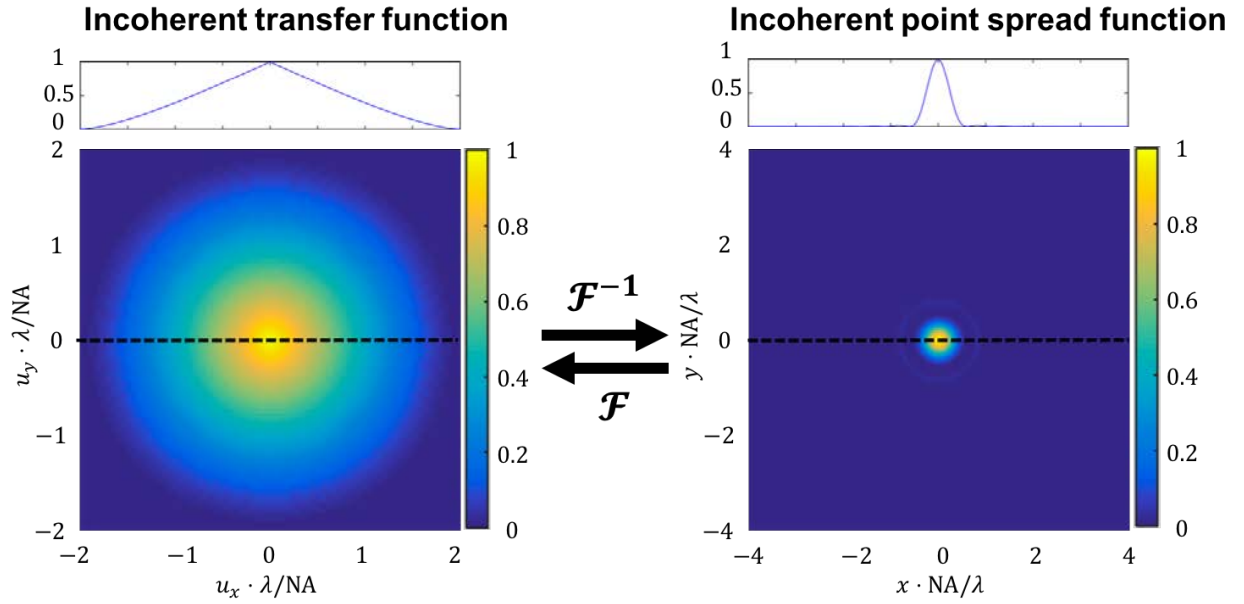


Figure 1.3: The 2D incoherent transfer function and its corresponding point spread function (PSF) with the profile plot showing the function value at the central cutline.

1.2 Super-resolution coherent imaging

In the coherent imaging scenario, we focus on observing light diffraction contrasts of samples. Since the diffraction-limited resolution originates from the limited collection angle of the diffracted light from the sample, the collection of higher-angle diffracted light is necessary to break the diffraction limit. Borrowing from the concept of synthetic aperture in the radar community, super-resolution techniques generally introduce ways to temporally multiplex higher-angle diffracted light into the aperture for super-resolution coherent imaging [2, 10, 11, 40–42].

The general strategy of synthetic aperture coherent imaging is to illuminate the sample with coherent plane waves at oblique angles. In terms of geometric optics, the diffractive angle of light is constant with respect to the angle of the illumination light. When we gradually increase the angle of oblique illumination, the angle of some originally detectable diffraction light becomes too large to detect. However, the angle of some previously not detectable diffraction light becomes small enough for the system to capture. Figure 1.4 demonstrates this phenomenon by comparing the scattered light using on-axis and oblique illumination. By varying the angle of oblique illumination and capturing multiple images, we are able to collect light from an overall larger range of angles.

In a more rigorous wave optics formulation, we write the output wave field under oblique illumination as

$$u_o(\mathbf{r}) = [o(\mathbf{r})e^{i2\pi\mathbf{v}_0 \cdot \mathbf{r}}] \otimes h_c(\mathbf{r}), \quad (1.14)$$

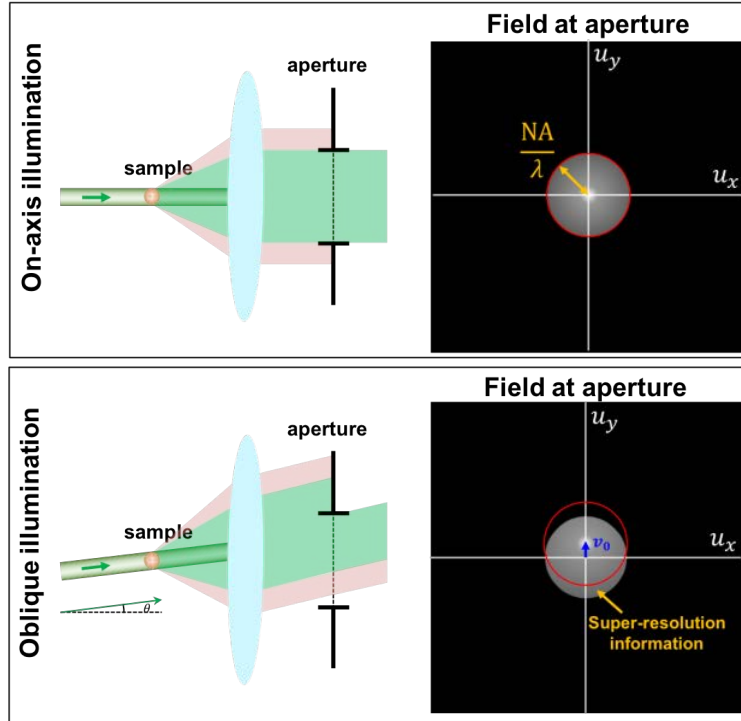


Figure 1.4: Comparison of scattered field at the aperture for on-axis illumination and oblique illumination. Coherent super-resolution information (higher-angle scattered light) is brought into the system by oblique illumination.

where we have dropped the time-varying notation and assumed that magnification $M = 1$ for simplicity. \mathbf{v}_0 is the spatial frequency of the illumination light, which is related to the angle of the illumination plane wave. For example, a x -tilted plane wave with angle, θ , has $\mathbf{v}_0 = (\sin \theta / \lambda, 0)$. By conducting a Fourier transform on this output field, we have

$$\tilde{u}_o(\mathbf{u}) = o(\mathbf{u} - \mathbf{v}_0) \cdot \tilde{h}_c(\mathbf{u}). \quad (1.15)$$

Thus, plane wave illumination on the sample shifts the Fourier space of the sample by \mathbf{v}_0 . This information then gets filtered by the coherent transfer function as shown in the field at aperture in Fig. 1.4. With proper selection of \mathbf{v}_0 , we can decide which part of the sample's Fourier space to transfer through the imaging system. Super-resolution coherent imaging is then implemented by combining multiple of these measurements computationally. The final achievable resolution depends on the final covered area of the Fourier space, which is set by the sum of the bandwidth of the transfer function NA/λ and the maximally achievable $\mathbf{v}_0 = NA_{\text{illum}}/\lambda$, where NA_{illum} is the illumination numerical aperture. The final resolution is roughly the reciprocal of the Fourier space bandwidth, which is $\lambda/(NA + NA_{\text{illum}})$.

The unmentioned problem here so far is that our sensor only measures intensity, instead of the electric field as in Eq. (1.8). A direct synthesis of these high-angle diffraction wave fields for super-resolution requires the detection of the complex wave field for linear super-

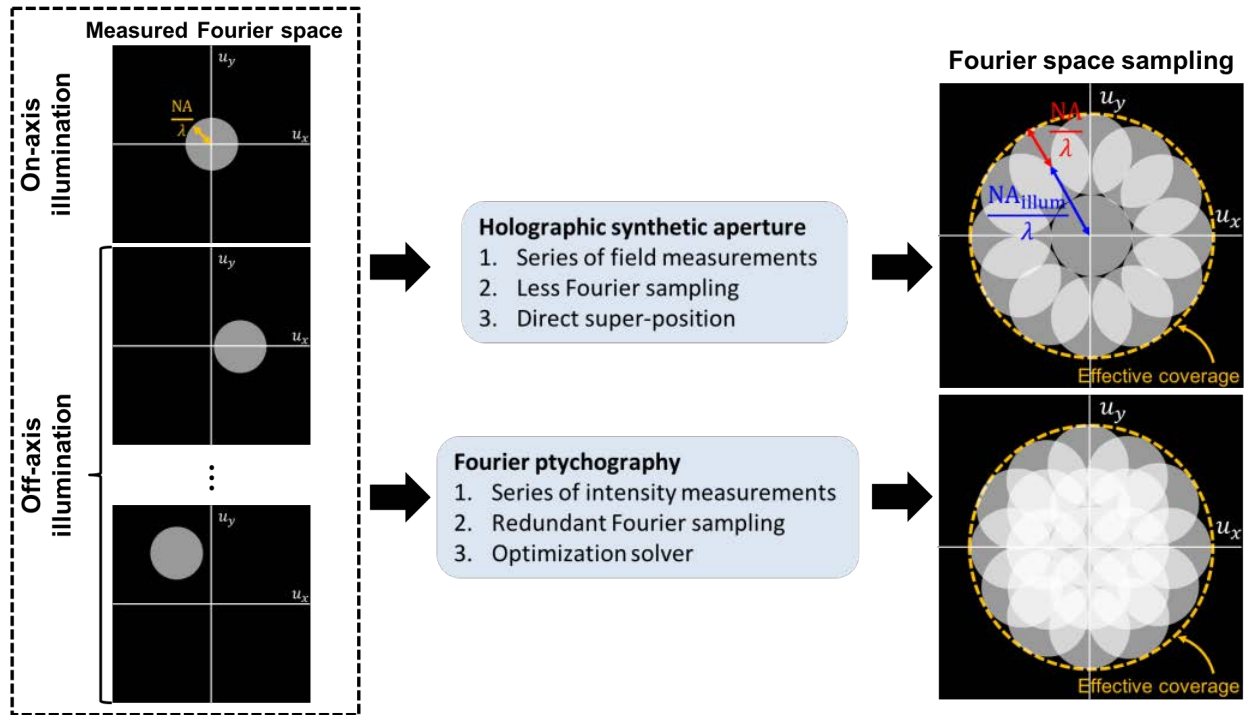


Figure 1.5: Comparison of holographic synthetic aperture microscopy and Fourier ptychographic microscopy in terms of Fourier space sampling, type of acquisition, and inversion solver.

position. An extra holography setup can achieve this type of detection. In fact, most of the synthetic aperture coherent imaging techniques are implemented with a holography setup equipped with oblique illumination capability [9, 11–16, 41]. An exception is a technique called Fourier ptychographic microscopy (FPM) proposed by [20]. FPM achieves multiple oblique illuminations efficiently with a LED array illuminator. A redundant sampling of \mathbf{v}_0 ensures a collection of high-angle diffraction intensity images with lots of overlapping parts in Fourier space. The super-resolution wave field can then be deduced with only these redundant intensity measurements using an optimization algorithm. Figure 1.5 summarizes the differences between holographic synthetic aperture microscopy and FPM. The holographic setup is efficient in terms of acquiring super-resolution information compared to FPM, and the inversion algorithm is relatively simpler. However, the FPM setup is simpler and much lower cost. In Chapter 2, we will go through a more detailed explanation of FPM and describe the performance of its inverse algorithms.

1.3 Super-resolution incoherent imaging

Previously, we have established the input-output relationship for the incoherent imaging scenario. The output measured intensity is a convolution between the intensity of the input

wave and the incoherent PSF. In the following discussion, we focus on one of the common incoherent imaging cases, fluorescence imaging, where we rewrite our imaging model as

$$I_o(\mathbf{r}) = I_i(\mathbf{r}) \otimes h_{inc}(\mathbf{r}), \quad (1.16)$$

where $I_i(\mathbf{r})$ is the input fluorescent intensity. When the illumination (excitation) intensity is uniform across the FOV, $I_i(\mathbf{r}) = o_f(\mathbf{r})$ is the fluorescent distribution of the sample at the input sample plane. Mathematically, we have the same problem as previously encountered in the coherent imaging case, where $o_f(\mathbf{r})$ is diffraction-limited. However, this time the linear space-invariant relationship is built upon the input-output intensity instead of the electric field. The previous approach to increase resolution of a coherent imaging system with oblique illumination gives uniform excitation intensity across the FOV, which results in the same output intensity measurement as normal uniform excitation. Thus, it does not encode super-resolution information. Different temporal multiplexing schemes are necessary in this case for super-resolution fluorescent information. Generally speaking, there are two main categories of methods that are able to encode super-resolution fluorescent information in separate measurements based on different physical mechanisms, illumination-based and single-molecule based methods.

1.3.1 Illumination based fluorescence super-resolution

In the coherent super-resolution case, varying plane wave illumination multiplexes information of different Fourier spaces into the aperture. Active illumination is key to super-resolution imaging. In fluorescence imaging, the strength of the fluorescent intensity scales linearly with the intensity of the excitation light (when excitation intensity is low and there is no saturation), so a plane wave modulation with any angle would result in constant excitation intensity map and thus no spatial modulation on the fluorescent intensity. Thus, varying illumination angles as in the coherent imaging case cannot be used. Instead of plane wave modulation, we can project intensity variation (structured illumination) on the fluorescent sample to create spatial modulation and multiplex super-resolution information into the aperture. This approach is called structured illumination microscopy (SIM) [5, 6].

With this spatial modulation, the input fluorescent intensity is

$$I_i(\mathbf{r}) = p(\mathbf{r}) \cdot o_f(\mathbf{r}), \quad (1.17)$$

where $p(\mathbf{r})$ is the illumination intensity pattern and is a real function. We further express the measured output intensity mathematically as

$$I_o(\mathbf{r}) = [p(\mathbf{r}) \cdot o_f(\mathbf{r})] \otimes h_{inc}(\mathbf{r}), \quad (1.18)$$

To understand the super-resolution capability, we observe the Fourier transform of Eq. (1.18) expressed as

$$\tilde{I}_o(\mathbf{u}) = [\tilde{p}(\mathbf{u}) \otimes \tilde{o}_f(\mathbf{u})] \cdot \tilde{h}_{inc}(\mathbf{u}) \quad (1.19)$$

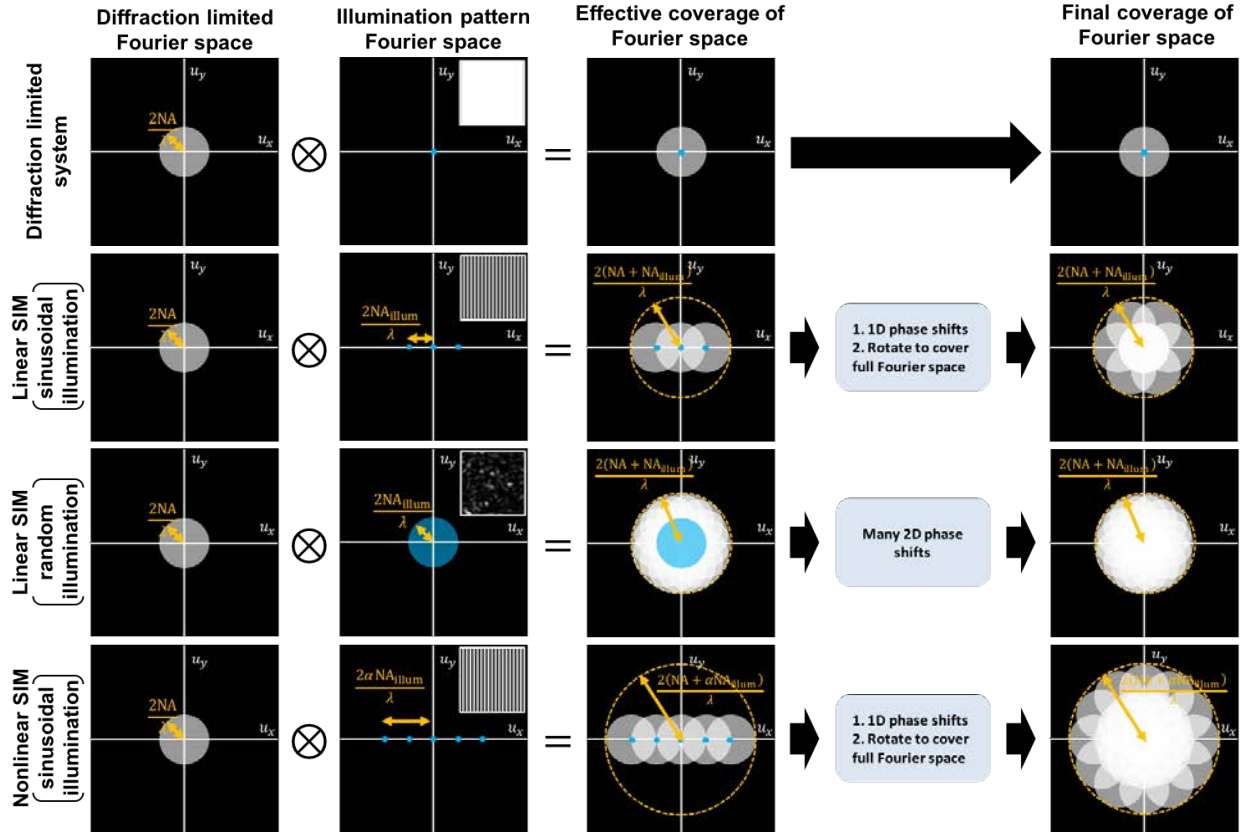


Figure 1.6: Fourier space coverage comparison for a diffraction-limited system, linear SIM under sinusoidal illumination, linear SIM under random illumination, and nonlinear SIM under sinusoidal illumination. The modulation between the structured pattern and the fluorescent sample in real space could be interpreted as a convolution of these variables' spectrum in the Fourier space. The nonlinear SIM is able to achieve larger Fourier space coverage because the nonlinearity helps to create pattern that is much larger than the illumination NA. The inscribed figures in the illumination pattern columns are the corresponding real space images of these illumination patterns.

For example, assuming the intensity variation is sinusoidal, the intensity modulation in the real and Fourier space are expressed as

$$\begin{aligned}
 p(\mathbf{r}) &= 1 + \cos(2\pi \mathbf{v}_0 \cdot \mathbf{r} + \phi) \\
 \tilde{p}(\mathbf{u}) &= \delta(\mathbf{u}) + \frac{e^{i\phi}}{2} \delta(\mathbf{u} - \mathbf{v}_0) + \frac{e^{-i\phi}}{2} \delta(\mathbf{u} + \mathbf{v}_0),
 \end{aligned} \tag{1.20}$$

where \mathbf{v}_0 and ϕ_0 is the spatial frequency and the phase of the illumination pattern. In this case, the Fourier space of the measured intensity is re-written as

$$\tilde{I}_o(\mathbf{u}) = \left[\tilde{o}_f(\mathbf{u}) + \frac{e^{i\phi}}{2} \tilde{o}_f(\mathbf{u} - \mathbf{v}_0) + \frac{e^{-i\phi}}{2} \tilde{o}_f(\mathbf{u} + \mathbf{v}_0) \right] \cdot \tilde{h}_{inc}(\mathbf{u}) \tag{1.21}$$

It is clear from Eq. (1.21) that multiple regions of Fourier space that were previously not detectable are multiplexed into this measurement. To de-multiplex this information, multiple measurements under the same illumination intensity but different phase shifts of the intensity pattern are required. Finally, these de-multiplexed regions of Fourier space are stitched together for super-resolution fluorescence imaging.

Generally speaking, the intensity modulation does not necessarily need to be sinusoidal pattern. As long as the Fourier space of the pattern is able to cover the targeted resolution, it can be any other pattern, such as a single focal spot (confocal microscope) [43, 44], multifocal spots [45–48] or even random speckle patterns [48–57], which will be adopted later in this thesis. Conditions with uniform, sinusoidal, and random illumination are illustrated in the first three rows of Fig. 1.6. We see how spatial modulation of different patterns result in larger coverage of Fourier space and the corresponding required shifts and rotations to fulfill the condition for super-resolution de-multiplexing. The final achievable resolution, similar to the coherent imaging case, depends on the final covered area of the Fourier space, which is set by the bandwidth of the transfer function $2NA/\lambda$ and the maximally achievable $\mathbf{v}_0 = 2NA_{\text{illum}}/\lambda$. The final resolution is roughly the reciprocal of the Fourier space bandwidth, which is $\lambda/2(NA + NA_{\text{illum}})$, suggesting a maximum of $2\times$ resolution gain for conventional epi-illumination fluorescence system.

However, this new SIM resolution limit is still breakable. In the previous paragraphs, we restricted ourselves to only discuss the case where the excitation intensity is weak enough to be linear with the fluorescent intensity. When the illumination intensity is stronger and the fluorescent rate starts to saturate, sinusoidal modulation becomes closer to square wave modulation. This is equivalent to applying a square wave intensity pattern, which contains a lot more high-frequency harmonics than the original illumination pattern, on the fluorescent sample (This can not be easily achieved with linear SIM because of limited illumination NA). Higher spatial frequency could be multiplexed into the system because of this, and thus resolution of $\lambda/2(NA + \alpha NA_{\text{illum}})$, where α is the number of the harmonics that are effective in bringing the super-resolution information. This is called nonlinear SIM [58], which utilizes fluorescent saturation, one of the fluorescent non-linear properties, for super-resolution. The last row of Fig. 1.6 summarizes how this saturation affects the Fourier space of the illumination pattern and then brings even more Fourier space information into the measurements. Other non-linear fluorescent properties (*e.g.* depletion, photo-switching, photo-activation, etc.) are also used together with structured illumination for resolution improvement [3, 59, 60].

For a conventional epi-illumination fluorescence system, $NA_{\text{illu}} = NA$, nonlinear SIM is the method to go beyond $2\times$ resolution gain previously set by standard linear SIM. In Chapter 4 and 5, we adopt transmission microscope geometry and demonstrate a method for greater than $2\times$ resolution gain with linear SIM, thus not requiring special fluorophores.

1.3.2 Single-molecule based fluorescence super-resolution

Another powerful approach to get super-resolution fluorescence imaging is based on identifying each single molecule with its statistical light emission behavior. The first two techniques proposed using this idea were photo-activated localization microscopy (PALM) [7] and statistical optical reconstruction microscopy (STORM) [8]. They both exploit the intrinsic light emission properties of fluorescent molecules to statistically turn on sparsely distributed fluorescent molecules for each exposure. Since the turned-on molecules are sparse, they can be successfully localized with a diffraction-limited microscope. By accumulating images of these statistically turned-on sparse sets of fluorescent molecules, a final super-resolution image with resolution up to single-molecule precision (usually on the order of 10nm) can be achieved. These methods have much to offer, but will not be further discussed in this thesis.

1.4 Computational super-resolution imaging

Until now, we have reviewed several major approaches to achieve super-resolution imaging for coherent and incoherent imaging systems. These previous techniques usually rely on an analytic mathematical relationship for direct information retrieval. For example, we see how measurements are linearly related to certain parts of the Fourier space in holography-based synthetic aperture microscopy (Eq.(1.15)) and linear SIM with sinusoidal illumination (Eq. (1.21)). This kind of relationship results in a simple one-step analytic inversion for super-resolution reconstruction, which is fast in computation, and easy for performance analysis.

However, analytic relationships cannot always be established. There are cases that enough information is captured for super-resolution reconstruction when there is no corresponding analytic inverse model for it. For example, in the case of Fourier ptychography, many scattering intensity measurements from angular illumination with redundant sampling is sufficient for a super-resolution phase image reconstruction, but there is no direct inversion relationship for this reconstruction. A similar situation applies to linear SIM with random illumination, PALM and STORM. Thanks to recent advances in convex optimization and compressed sensing [61, 62], the rekindled enthusiasm for optimization in the optical imaging community has triggered advances in many imaging problems, demonstrating the power of its general inversion framework. With optimization, problems like FPM, random-illumination SIM, PALM, and STORM are nicely formulated and solved. Though this type of inversion approach is usually associated with instability due to model mismatch, slow computation due to iterative algorithms, and difficulty of information analysis due to non-linearity, it provides a more general framework of information retrieval and correspondingly a more flexible design of optics to get the information.

Hence, incorporating the optimization component into the imaging problem potentially leads to different trade-offs in the optical design process, which helps spawn new application spaces. This approach is called computational imaging. Successful examples include

high-content imaging techniques [18–22, 30], where people break traditional sensor pixel-limited resolution for lensless microscopy [18, 19] and achieve more efficient resolution tiling microscopy over traditional FOV tiling microscopy [20–22, 30] with proper formulation of optimization to target high-throughput imaging applications. Another example is imaging techniques in extreme imaging systems (X-ray, TEM) with limited optical design degrees of freedom [63–67], where people have to exploit the generality of optimization formulation to accommodate physically feasible information capturing schemes for efficient information retrieval. The generality of the optimization formulation also provides a platform for people to easily include prior knowledge of their targeted unknown variable into the inversion process, which reduces the number of measurements required for proper information retrieval [62, 68, 69]. These types of compressed sensing approaches toward imaging are also part of computational imaging, but I will not talk too much about it in this thesis.

In this thesis, we first investigate existing computational imaging techniques for coherent and incoherent super-resolution imaging such as FPM and SIM with random illumination. Detailed optimization analysis of FPM leads us to discover the important role of the noise model in an optimization for computational imaging techniques. Proper selection of the noise model provides better robustness of information retrieval. Along with this investigation of FPM and SIM, we also modeled experimental errors or unknown experimental parameters such as system aberration, light source misalignment, and unknown random illumination patterns into the inversion process to help self-calibrate our model for better-quality reconstruction. In the end, we adopt this computational imaging approach, where the optimization formulation is incorporated into the process of optical design, and come up with a speckle illumination based imaging technique with new trade-offs for joint fluorescence and phase super-resolution microscopy.

1.5 Outline of thesis

The arrangement of this thesis shows progressive building blocks towards the final computational framework for joint fluorescence and phase super-resolution microscopy with speckle illumination. The summary of each chapter is described in the following points.

- Chapter 2 introduces an optimization algorithm analysis of a computational super-resolution technique for coherent imaging, Fourier ptychographic microscopy (FPM). FPM is a phase imaging technique that targets at the large fields-of-view and high-resolution biological imaging. With a replacement of an LED array on a standard commercial microscope, multi-angle illumination, which encodes multi-fold super-resolution phase information, can be easily implemented with high speed. Those super-resolution encoded measurements are then processed with a nonlinear optimization algorithm for super-resolution phase reconstruction. Our analysis focuses on reviewing and comparing multiple optimization algorithms of this problem using simulation and experimental data, revealing the surprising impact of the noise model adopted in different algorithms

on final reconstructions. It also explores additional self-calibration algorithms for better correcting experimental errors that cause model mismatch. Overall, this work led to a more experimentally robust algorithm for FPM.

- Chapter 3 proposes a more robust optimization-assisted algorithm for structured illumination fluorescence microscopy with random illumination patterns. Standard structured illumination microscopy uses organized periodic patterns for super-resolution encoding. High quality super-resolution reconstruction depends on accurate knowledge of the illumination pattern, which requires well-aligned and aberration-free optics. Random speckle illumination for super-resolution encoding is relatively easy to achieve and maintain. However, the inverse problem becomes more complicated and unstable. In this work, we studied better uses of the statistical information of the random patterns and our proposed method out-performed other contemporary algorithms for the same purpose, achieving better super-resolution fluorescence imaging with random patterns.
- Chapter 4 leverages the experiences from the previous chapters and proposes a combined framework for super-resolution phase and fluorescence multimodal microscopy using scanned unknown speckle illumination. We use Scotch tape as a patterning element to efficiently generate large-area speckle illumination that has much higher illumination angle than the detectable angle of a low-NA objective lens. With proper optimization and self-calibration formulation, we are able to achieve $4\times$ resolution improvement, resulting in sub-cellular resolution imaging over millimeter scale FOV for both fluorescence and phase contrast, which leads us to high-content imaging applications. Computational imaging techniques for high-content imaging usually provide a single contrast. We demonstrate that computational imaging plus Scotch-tape-enabled speckle illumination is an efficient solution for multimodal high-content imaging.
- Chapter 5 generalizes the theoretical framework from Chapter 4 to a multiple scattering 3D model to enable 3D super-resolution fluorescence and phase imaging with unknown speckle illumination, accommodating situations of thick samples and shallow depth-of-field of speckles. This provides a potential solution for 3D multimodal high-content imaging.
- Chapter 6 states conclusions and future work.

Chapter 2

Optimization analysis for Fourier ptychographic phase microscopy

In Chapter 1, we have reviewed how super-resolution imaging is achieved in coherent imaging. The general approach is to illuminate the sample with multiple different angles to collect high-angle diffracted light, which contains high-resolution information of the scattering sample. Among many examples of coherent super-resolution imaging techniques, Fourier ptychographic microscopy (FPM) is a computational imaging approach (optimization-based solver) that does not require interferometric measurements. By capturing a stack of low-resolution images under dense sampling of different illumination angles, an inverse optimization algorithm can be used to computationally reconstruct the high-resolution complex field.

As a stepping stone for developing further computational imaging methods in different imaging scenarios, we first choose FPM as an computational imaging example to study the concept of coherent super-resolution imaging and the optimization algorithm for image reconstruction. In this chapter, we explain the concept of FPM and how experimentally we implement this technique with an LED array. With this basic understanding, we focus our study on the inverse optimization algorithm of FPM and then compare multiple proposed algorithms in terms of experimental robustness.

In this study, we find that the main sources of error are noise, aberrations and mis-calibration (i.e. *model mis-match*). Using simulations and experiments, we demonstrate that the choice of cost function plays a critical role, with *amplitude-based* cost functions performing better than *intensity-based* ones. The reason for this is that Fourier ptychography datasets consist of images from both brightfield and darkfield illumination, representing a large range of measured intensities. Both noise (e.g. Poisson noise) and model mis-match errors are shown to scale with intensity. Hence, algorithms that use an appropriate cost function will be more tolerant to both noise and model mis-match. Given these insights, we propose a global Newton's method algorithm which is robust and accurate. Finally, we discuss the impact of procedures for algorithmic correction of aberrations and mis-calibration.

2.1 Introduction

2.1.1 LED array microscope

The LED array microscope is a powerful platform for computational coherent microscopy in which a wide range of capabilities are enabled by a single hardware modification to a traditional brightfield microscope - the replacement of the source with a programmable LED array as shown in Fig. 2.1(a) [70]. Each LED in the array corresponds to illumination of the sample by a unique angle. In Fourier optics, illuminating the sample with different angles is equivalent to shifting the sample's Fourier space in different amounts as in Eq. (1.15) shows. Thus, this simple, inexpensive hardware modification allows programming of the Fourier space measurement of our sample.

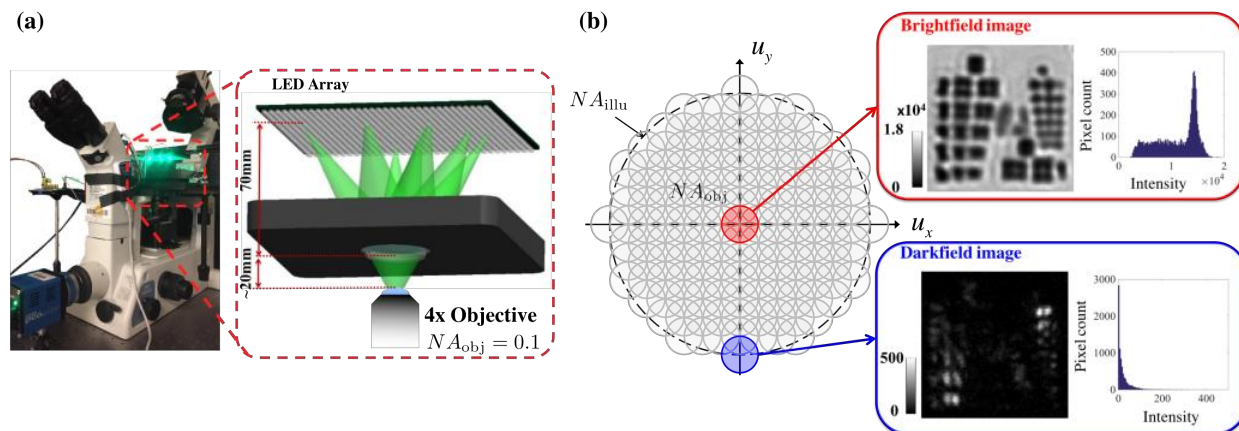


Figure 2.1: (a) Experimental setup for Fourier ptychography with an LED array microscope. (b) The sample's Fourier space is synthetically enlarged by capturing multiple images from different illumination angles. Each circle represents the spatial frequency coverage of the image captured by single-LED illumination. Brightfield images have orders of magnitude higher intensity than darkfield (see histograms), resulting in different noise levels.

The range of illumination angles that can be programmed is much larger than the range of angles that pass through the objective [set by its numerical aperture (NA_{obj})]. This means that illumination by the central LEDs produces brightfield images, whereas illumination by the outer LEDs (outside the NA_{obj}) produces dark field images as shown in Fig. 2.1(b) [70]. Alternatively, by sequentially taking a pair of images with either half of the source on, we obtain phase derivative measurements by differential phase contrast (DPC) [71–76]. Finally, a full sequential scan of the 2D array of LEDs (angles), while taking 2D images at each angle, captures a 4D dataset similar to a light field [77] or phase space measurement [78]. This enables all the computational processing of light field imaging. For example, angular information can be traded for depth by using digital refocusing algorithms to get 3D intensity or 3D phase contrast [79].

Up to this point, all these techniques can be realized by a simple replacement of the light source with an LED array. The capability of combining digital signal processing and optical imaging has been well demonstrated with the LED microscope. In the following, we will talk about one more powerful capability of this microscope: Fourier ptychography.

2.1.2 Fourier ptychography

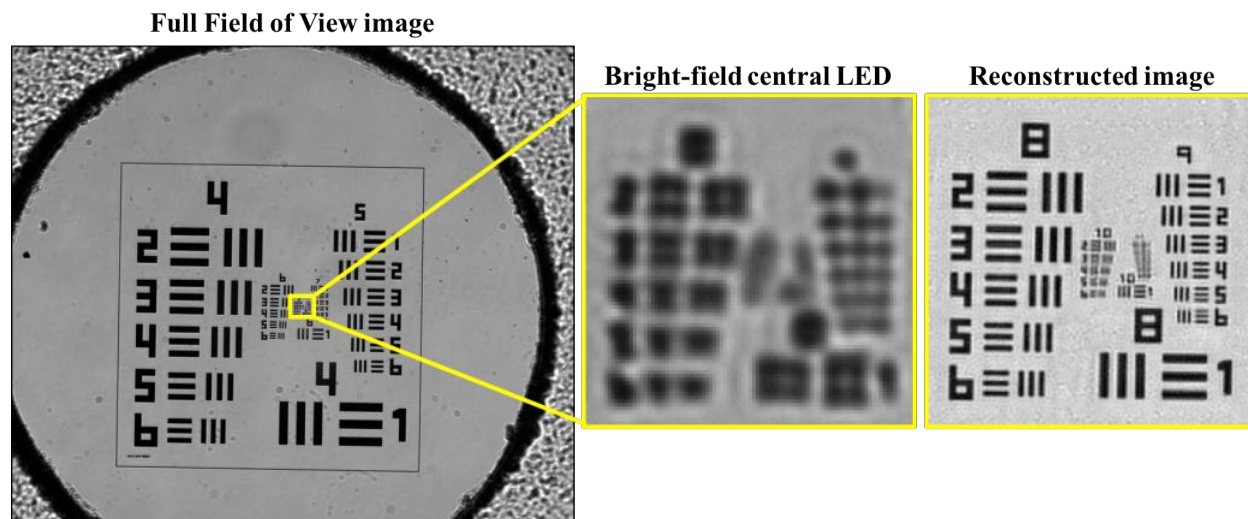


Figure 2.2: The full field of view and the zoomed region of the USAF resolution target for the brightfield raw image from central LED illumination and FPM reconstructed image in the central region of the FOV, demonstrating the super-resolution capability.

In the previous paragraph and chapter, taking images while sequentially scanning LEDs in the array corresponds to measuring the Fourier space of the sample in different regions as Fig. 2.1(b) and Eq. (1.15) show. Brightfield and darkfield measurements could be obtained when the illumination angle is smaller and larger than the objective NA (NA_{obj}), respectively. Normally, a holography setup is necessary to measure the scattered electric field at each illumination angle for further super-resolution synthesis from direct superposition as discussed in Chapter 1. However, by densely sampling the scanned angle for redundant Fourier space measurements, a nonlinear optimization algorithm could directly synthesize these measurements without interferometric detection and enable a super-resolution complex field reconstruction with effective NA of the sum of illumination and objective NA ($NA_{eff} = NA_{illu} + NA_{obj}$). This approach is essentially a combination of synthetic aperture and translational-diversity phase retrieval [63, 80], which is termed Fourier Ptychographic microscopy (FPM) [20]. Since illumination NA could easily go a lot higher than the NA of a low-NA objective with the LED array microscope, FPM is usually implemented on a low-NA microscope system to enable large field of view (FOV) and high resolution across the entire image, resulting in gigapixel images as shown in Fig. 2.2. With access to super-resolution

quantitative phase contrast using FPM, we can image biological cells with better contrast than a traditional brightfield microscope as shown in Fig. 2.3. The phase image contains quantitative shape and density information about the samples, providing details about small sub-cellular structures.

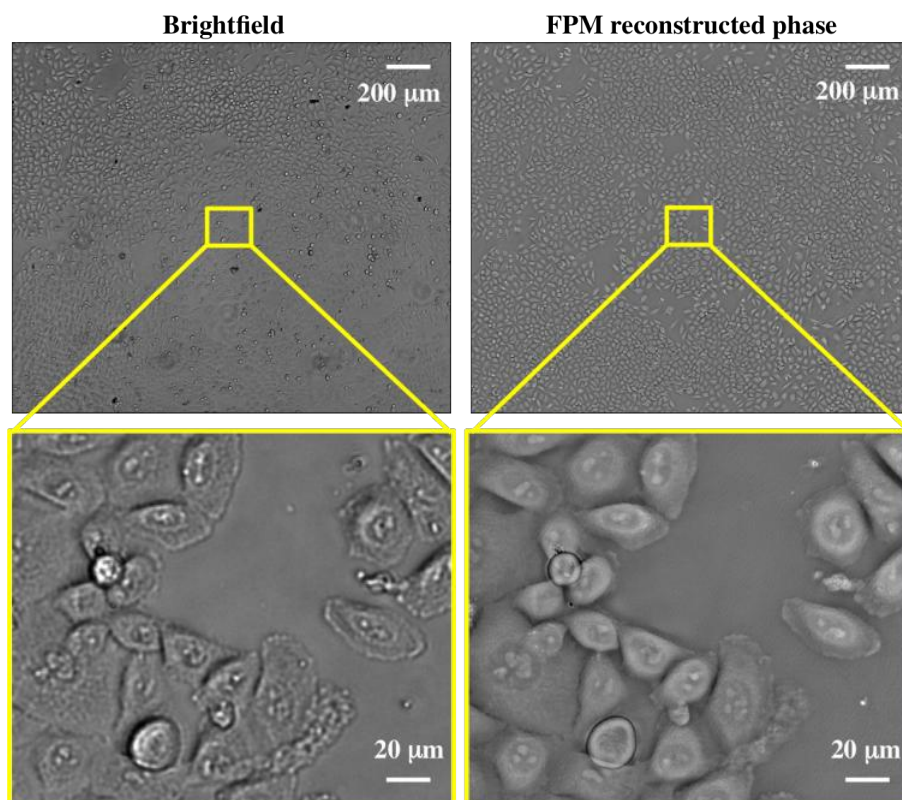


Figure 2.3: The full field of view and the zoomed region for both brightfield and FPM reconstructed phase of the HeLa cell, respectively.

On the algorithm side, FPM is a large-scale optimization problem. A single full FOV image taken under microscope contains $\sim 2000 \times 2000$ pixels and around 300 images are taken in each experiment. Since it is a super-resolution phase imaging technique, the final reconstruction has $\sim 10^4 \times 10^4$ pixels of complex values to solve for, which corresponds to 0.2 giga-variables. In addition, this problem is a phase retrieval problem and thus a nonlinear and non-convex optimization problem. It is then challenging to prove algorithm convergence to the global minimum. However, many people have shown successful reconstruction consistently in practice when appropriate parameters are used. This is one of the important issues that will be discussed in this chapter.

On the application side, FPM's scan-free high space bandwidth product imaging capability has great potential for revolutionizing biomedical imaging, with applications in optical disease diagnosis, digital pathology [20, 81–83] and *in vitro* live cell imaging [22]. The origi-

nal FPM method only applies to 2D thin objects, however, new models and reconstruction algorithms also enable 3D reconstruction of thick samples [79]. The ability to achieve such capabilities with a simple and inexpensive hardware modification to a commercial microscope (with no moving parts) opens up new opportunities for opensource distribution and wide-scale adoption in biological imaging applications requiring large-scale in vitro microscopy.

2.1.3 Overview of Fourier ptychographic phase retrieval algorithms

Multiple algorithms have been proposed for solving the nonlinear non-convex inverse FPM problem, which amounts to phase retrieval and synthetic aperture. Amongst these, there are the usual trade-offs between accuracy, noise performance and computational complexity. Here, we review the first FPM algorithm and several other later proposed optimization algorithms for comparison. We also classify these algorithms based on their gradient orders, convergence claims, and implementations. In the end, we summarize our findings and contribution to FPM algorithms from this study.

2.1.3.1 Gerchberg-Saxton FPM algorithm

The original FPM algorithm used a Gerchberg-Saxton approach [84], which is a type of alternating projections [85–88], first developed for traditional ptychography [63, 80, 89–92] and later for FPM [20, 21, 93]. Shifted support constraints (finite pupil size) are enforced in the Fourier domain as the corresponding amplitude constraints (measured images) are applied in the image domain, while letting the phase evolve as each image is stepped through sequentially. The Gerchberg-Saxton method, which is a type of gradient descent, represents a natural way to solve phase retrieval problems by trying to directly minimize some cost function that describes the differences between actual and predicted measurements. Unfortunately, these formulations are often non-convex in nature and do not come with global convergence guarantees.

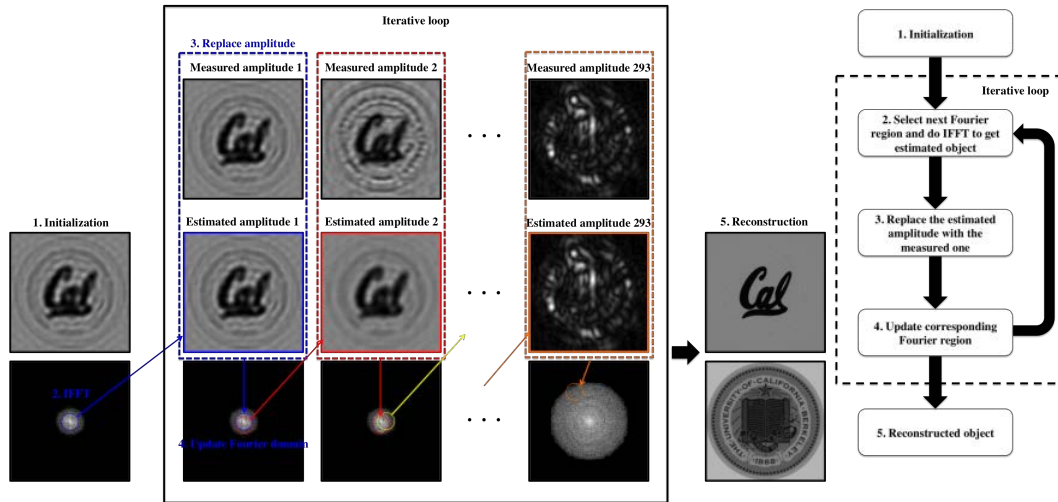


Figure 2.4: Schematic of the phase retrieval procedure for the Gerchberg-Saxton (gradient descent) iterative algorithm. The sample’s complex field is iteratively updated as its estimate is propagated back and forth between real space and Fourier space constraints, for each of the measured intensity images taken at multiple Fourier regions.

Since this is the first proposed algorithm to solve FPM, we give a brief introduction on how the algorithm stitch all low-resolution image to get one high-resolution phase image. Figure 2.4 shows the reconstruction procedure for Gerchberg-Saxton approach. First, we use the on-axis illuminated image as our initialization. Second, we crop the Fourier region corresponding to the image illuminated with the next LED and then do inverse Fourier transform to get our estimate complex field. Third, we replace the estimated amplitude with the measured amplitude from our dataset. Fourth, we Fourier transform the updated image and put this information to its corresponding region. By doing this process through all the images and repeating this whole process many times, we can get a final reconstruction as shown in the fifth step of Fig. 2.4.

2.1.3.2 Other FPM algorithms

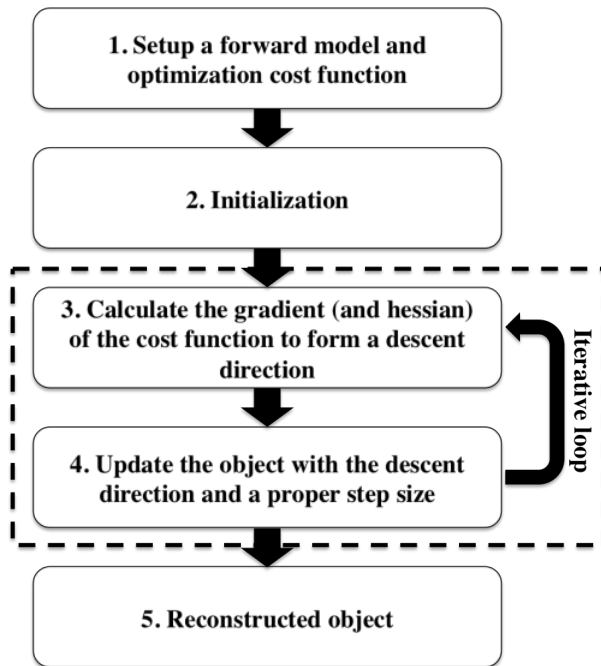


Figure 2.5: The general flow of iterative nonlinear optimization algorithms for Fourier ptychographic microscopy.

In general, the procedures of all proposed Fourier ptychographic phase retrieval algorithms including the Gerchberg-Saxton method can be summarized by an optimization procedure as shown in Fig. 2.5. They are all based on minimizing a certain cost function defined by measuring the difference between the measured intensity images and the estimated intensity from propagating the estimate through the forward model. We first give an initialization as mentioned in the last paragraph. To minimize the cost function, we then follow an iterative process, which involves calculating the gradient (and possibly Hessian) of the cost function. Using the gradient (and possibly Hessian), we form the search direction and update the estimation along this direction with appropriately chosen step size. We then repeat this process until we reach convergence. In the following, we will give an overview of the recently proposed algorithm in the context of optimization and put them into different classes. A brief introduction of algorithms that have been proposed to solve the FPM problem:

- **Gradient descent:** Use the gradient of the cost function and a proper step size to minimize the cost function value to update the unknown variables.
- **Wirtinger Flow:** Similar to gradient descent but with special initialization and step size to guarantee global convergence of the coded-mask phase retrieval problem [94].

- **Newton’s method:** Approximate the cost function to be a quadratic function and use both the gradient and Hessian of the cost function to give a better update direction of the unknown variables. If the Hessian is approximated, the method is called a quasi-Newton method. Gauss-Newton method is a kind of quasi-Newton method, which approximates the Hessian with the outer product of the gradient.
- **PhaseLift:** Instead of solving the known variable vector \mathbf{x} , this method reformulates the phase retrieval problem and solve for optimal \mathbf{xx}^H . For this case, the problem becomes convex and there will be a global convergence, at a cost of significantly increased computational complexity.

2.1.3.3 Classification of FPM algorithms

Depending on whether we only use the gradient or we use both the gradient and Hessian to update our complex field, the algorithm can be classified as a *first-order* method or a *second-order* method, respectively. Gradient descent and Wirtinger Flow [94] are *first-order* methods. The Wirtinger Flow algorithm has been shown to have global convergence guarantees. This method has been successfully applied to FPM [95], though the actual implementation deviates from theory somewhat. In the Wirtinger Flow framework, the optimization procedure is similar to gradient descent, except that the step size and initial guess are carefully chosen for provable convergence. For the *second-order* methods (e.g. Newton’s method), they have been shown to provide faster convergence rates [96]. While computing second-order derivatives increases complexity, we find that it usually reduces the number of iterations needed, enabling fast overall run times.

Another class of algorithms that have been proposed are based on convex relaxations [97–101]. This class of phase retrieval algorithms, called PhaseLift, re-frames the problem in higher dimensions such that it becomes convex, then aims to minimize the cost function between actual and predicted intensity via semidefinite programming. These algorithms come with the significant advantage of rigorous mathematical guarantees [102] and were successfully applied to FPM data [101]. The actual implementations of these algorithms, however, deviate from the provable case due to computational limitations.

Algorithms can be further classified as sequential or global, depending on whether the update is done for each image, one at a time (sequentially), or all at once with the full set of images (globally) for each iteration. Global methods are expected to perform better, at a cost of additional computational requirements. In our studies, results show little difference between the sequential and global implementation of any particular algorithm (see Fig. 2.9), suggesting that sequential procedures may be sufficient, allowing reduced computational requirements.

One seemingly unimportant classification of algorithms is whether their cost function minimizes differences in *intensity* or *amplitude*. Throughout this chapter, we refer to algorithms that minimize intensity differences as *intensity-based* algorithms, and algorithms that minimize amplitude differences as *amplitude-based* algorithms. Since intensity is amplitude

squared, both drive the optimization in the correct direction; hence, one might expect that the choice between the two is of little consequence. **Surprisingly, however, we find that the cost function is the key predictor of experimental performance for our experimental dataset.** *Intensity-based* algorithms suffer from strong artifacts (see Fig. 2.9), which we show to be due to noise and model mis-match errors. Hence, amplitude-based algorithms perform better on imperfect data, so are more robust. Our goal is to explain why this happens.

2.1.3.4 Contribution of this chapter

We will show that in order for a phase retrieval scheme to be robust to experimental imperfections, the choice of cost function is of crucial importance. One source of error in our experimental data is measurement noise, including Gaussian noise or Poisson shot noise. Another main source of error is model mis-match, caused by experimental imperfections such as aberrations and LED misalignment. A particular problem of FPM datasets is that they contain both brightfield and darkfield images, which have drastically different intensity levels (see Fig. 2.1). Brightfield images can have several orders of magnitude higher intensity than darkfield images; thus, the amount of Poisson noise will also be significantly higher. If this difference in the noise levels is not properly accounted for, the brightfield noise may drown out the darkfield signal. We will further show that aberrations and LED mis-calibration - the two main model mis-match errors in our experiments - result also in intensity-dependent errors. Thus, by carefully designing the the cost function, we can develop algorithms that are significantly more robust to *both* noise and model mis-match.

In Section 2.2 and 2.3, we develop a maximum likelihood theory which provides a flexible framework for formulating the FPM optimization problem with various noise models. In particular, we will focus on Gaussian and Poisson noise models. We find that *amplitude-based* algorithms effectively use a Poisson noise model, while *intensity-based* algorithms use a Gaussian noise model. In Section 2.4, we simulate four FPM datasets, three of which are contaminated with measurement errors (see Fig. 2.6): Poisson noise, aberrations, and LED misalignment. We compare the performance of various algorithms on these datasets to demonstrate that the imperfections in our experimental data are more consistent with a Poisson noise model. This explains our observations that *amplitude-based* algorithms are more experimentally robust than *intensity-based* algorithms. In Section 2.5, we develop additional step in original FPM algorithm to correct for the two main model mis-matches, which are aberrations and LED mis-calibration.

2.2 Optimization formulation to Fourier ptychographic microscopy

After providing an introduction to Fourier ptychographic microscopy and its algorithms, here we will present a more detailed formulation. In this section, we start from introducing

the physical forward model of Fourier ptychography and then move on to talk about how to use this forward model to form optimization problems. Based on noise assumptions with different statistics, the cost functions of the FPM optimization problem can be classified as intensity-based (white Gaussian noise), amplitude-based (most commonly used), and the Poisson-likelihood-based (Poisson noise) cost functions. These cost functions are then used in later sections to derive different FPM algorithms. In the end of this section, we define the notation used in this thesis.

2.2.1 Forward problem for Fourier ptychography

Consider a thin sample with transmission function $o(\mathbf{r})$, where $\mathbf{r} = (x, y)$ represents the 2D spatial coordinates in the sample plane. Assuming that the LED array is sufficiently far from the sample, each LED will illuminate the sample by a plane wave from a different angle, defined by $\exp(i2\pi\mathbf{u}_\ell \cdot \mathbf{r})$, where $\mathbf{u}_\ell = (u_{\ell,x}, u_{\ell,y})$ is the spatial frequency corresponding to the ℓ -th LED, $\ell = 1, \dots, N_{\text{img}}$. After passing through the sample, the exit wave is the product of the sample and illumination complex fields, $o(\mathbf{r}) \exp(i2\pi\mathbf{u}_\ell \cdot \mathbf{r})$. The tilted plane wave illumination means that the Fourier transform of this exit wave is just a shifted version of the Fourier spectrum of the object, $O(\mathbf{u} - \mathbf{u}_\ell)$, where $O(\mathbf{u}) = \mathcal{F}\{o(\mathbf{r})\}$ and \mathcal{F} is the 2D Fourier transform. This exit wave then passes through the objective lens, where it is low-pass filtered by the pupil function, $P(\mathbf{u})$, which is usually a circle with its size defined by NA_{obj} . Finally, with \mathcal{F}^{-1} being the 2D inverse Fourier transform, we can write the intensity at the image plane as [21]

$$I_\ell(\mathbf{r}) = |\mathcal{F}^{-1}\{P(\mathbf{u})O(\mathbf{u} - \mathbf{u}_\ell)\}|^2. \quad (2.1)$$

2.2.2 Optimization problem based on different noise models

Most algorithms solve the FPM problem by minimizing the difference between the measured and estimated amplitude (or intensity), without assuming a noise model. Hence, the FPM problem can be formulated as the following optimization

$$\min_{O(\mathbf{u})} f_A(O(\mathbf{u})) = \min_{O(\mathbf{u})} \sum_{\ell} \sum_{\mathbf{r}} |\sqrt{I_\ell(\mathbf{r})} - \mathcal{F}^{-1}\{P(\mathbf{u})O(\mathbf{u} - \mathbf{u}_\ell)\}|^2. \quad (2.2)$$

Since the cost function here, $f_A(O(\mathbf{u}))$, aims to minimize the difference between the estimated amplitude and the measured amplitude, this is the amplitude-based cost function. By optimizing this cost function, the projection-based algorithms for Fourier ptychography can be obtained [20, 21, 93], which treat each measurement as an amplitude-based sub-optimization problem. This formulation is used in the traditional Gerchberg-Saxton phase retrieval approach.

If we have information about the statistics of the noise, we can use it in our optimization formulation via the maximum likelihood estimation framework [64]. If we assume that our measured images suffer only from white Gaussian noise, then the probability of capturing

the measured intensity $I_\ell(\mathbf{r})$ at each pixel, given the estimate of $O(\mathbf{u})$, can be expressed as

$$p[I_\ell(\mathbf{r})|O(\mathbf{u})] = \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left[-\frac{(I_\ell(\mathbf{r}) - \hat{I}_\ell(\mathbf{r}))^2}{2\sigma_w^2}\right], \quad (2.3)$$

where $\hat{I}_\ell(\mathbf{r}) = |\mathcal{F}^{-1}\{P(\mathbf{u})O(\mathbf{u} - \mathbf{u}_\ell)\}|^2$ and σ_w is the standard deviation of the Gaussian noise. $\hat{I}_\ell(\mathbf{r})$ and $I_\ell(\mathbf{r})$ denote the estimated and measured intensity, respectively.

The likelihood function is the overall probability due to all the pixels in all the images and can be calculated as $\prod_\ell \prod_{\mathbf{r}} p[I_\ell(\mathbf{r})|O(\mathbf{u})]$, assuming measurements from all pixels are independent. In maximum likelihood estimation, the goal is to maximize the likelihood function. However, it is easier to solve this problem by turning the likelihood function into a negative log-likelihood function which can be minimized. The negative log-likelihood function associated with this probability distribution can be calculated as

$$\begin{aligned} \mathcal{L}_{\text{Gaussian}}(O(\mathbf{u})) &= -\log \prod_\ell \prod_{\mathbf{r}} p[I_\ell(\mathbf{r})|O(\mathbf{u})] \\ &= \sum_\ell \sum_{\mathbf{r}} \left[\frac{1}{2} \log(2\pi\sigma_w^2) + \frac{(I_\ell(\mathbf{r}) - \hat{I}_\ell(\mathbf{r}))^2}{2\sigma_w^2} \right]. \end{aligned} \quad (2.4)$$

The next step is to minimize this negative log-likelihood function by estimating $O(\mathbf{u})$ so that the overall probability is maximized. For white Gaussian noise, it is assumed that σ_w^2 are the same across all pixels for all images (i.e. all measurements have the same amount of noise), though this will *not* be the case for FPM datasets. By making a Gaussian noise assumption, the first term in (2.4) is a constant and can be ignored. The optimization problem then reduces to

$$\min_{O(\mathbf{u})} f_I(O(\mathbf{u})) = \min_{O(\mathbf{u})} \sum_\ell \sum_{\mathbf{r}} |I_\ell(\mathbf{r}) - |\mathcal{F}^{-1}\{P(\mathbf{u})O(\mathbf{u} - \mathbf{u}_\ell)\}|^2|^2. \quad (2.5)$$

We call this cost function, $f_I(O(\mathbf{u}))$, the intensity-based cost function because it aims to minimize the difference between the estimated intensity and the measured intensity. It also implies that noise from each pixel is treated the same and independent of the measured intensity. It will be shown later that the previous implementations of PhaseLift [101] and Wirtinger flow algorithms [95] for FPM aimed to optimize this intensity-based cost function. However, both can be implemented instead with a Poisson likelihood cost function.

If we assume instead that our measured images suffer from Poisson shot noise, then the probability of the measured intensity, $I_\ell(\mathbf{r})$, given the estimate of $O(\mathbf{u})$ can be expressed as

$$p[I_\ell(\mathbf{r})|O(\mathbf{u})] = \frac{[\hat{I}_\ell(\mathbf{r})]^{I_\ell(\mathbf{r})} \exp[-\hat{I}_\ell(\mathbf{r})]}{I_\ell(\mathbf{r})!} \approx \frac{1}{\sqrt{2\pi\sigma_{\ell,\mathbf{r}}^2}} \exp\left[-\frac{(I_\ell(\mathbf{r}) - \hat{I}_\ell(\mathbf{r}))^2}{2\sigma_{\ell,\mathbf{r}}^2}\right]. \quad (2.6)$$

Note that the Poisson distribution is used to describe the statistics of the incoming photons at each pixel, which is a discrete probability distribution. Here, we assume that the intensity

is proportional to the photon count, so we can treat the distribution of the intensity as a Poisson distribution. When the expected value of the Poisson distribution is large, then this Poisson distribution will become more like a Gaussian distribution having a standard deviation proportional to the square root of the intensity, $\sigma_{\ell, \mathbf{r}} \approx \sqrt{I_{\ell}(\mathbf{r})}$, from the central limit theorem. This means that a large measured intensity at a particular pixel will imply large noise at that pixel. In the simulation, we impose Poisson noise on the measured intensity by distributing each pixel value with a Gaussian distribution and setting the standard deviation to $20\sqrt{I_{\ell}(\mathbf{r})}$. The negative log-likelihood of the Poisson noise model can then be calculated; the optimization problem is formed by minimizing the negative log-likelihood function with estimation of $O(\mathbf{u})$,

$$\begin{aligned} \min_{O(\mathbf{u})} \mathcal{L}_{\text{Poisson}}(O(\mathbf{u})) &= \min_{O(\mathbf{u})} \sum_{\ell} \sum_{\mathbf{r}} (-I_{\ell}(\mathbf{r}) \log[\hat{I}_{\ell}(\mathbf{r})] + \hat{I}_{\ell}(\mathbf{r}) + \log[I_{\ell}(\mathbf{r})!]) \\ &\approx \min_{O(\mathbf{u})} \sum_{\ell} \sum_{\mathbf{r}} \frac{(I_{\ell}(\mathbf{r}) - \hat{I}_{\ell}(\mathbf{r}))^2}{2\sigma_{\ell, \mathbf{r}}^2}. \end{aligned} \quad (2.7)$$

This cost function comes from the likelihood function of the Poisson distribution, so we call it the Poisson-likelihood-based cost function. It implies that the pixels with larger measured intensity are weighted smaller because they suffer from more noise. Since the brightfield images have more large-value pixels, they are assumed to be more noisy and thus are weighted smaller in the cost function. It is shown in the next section that the gradient of this cost function (2.19) is very similar to that of the amplitude-based cost function (2.16), which suggests that the amplitude-based cost function deals well with Poisson-like noise or model mis-match.

2.2.3 Vectorization Notation

For multivariate optimization problems such as (2.2) and (2.5), it is convenient to reformulate the problem using linear algebra. First, the functions need to be vectorized. Each of the captured images, $I_{\ell}(\mathbf{r})$, having $m \times m$ pixels, are raster-scanned into vectors, \mathbf{I}_{ℓ} , with size $m^2 \times 1$. Since the estimated object transmission function will have higher space-bandwidth product than the raw images, the estimated object should have $n \times n$ pixels, where $n > m$. For convenience, we actually solve for the Fourier space of the object, $O(\mathbf{u})$, which is vectorized into a vector \mathbf{O} with size $n^2 \times 1$. Before multiplying the pupil function, the Fourier space of the object is downsampled by a $m^2 \times n^2$ matrix \mathbf{Q}_{ℓ} . The matrix \mathbf{Q}_{ℓ} transforms a $n^2 \times 1$ vector into a $m^2 \times 1$ vector by selecting values out of the original vector, so the entries of this matrix are either 1 or 0 and each row contains at most one nonzero element. The pupil function $P(\mathbf{u})$ is vectorized into a vector \mathbf{P} with size $m^2 \times 1$. The 2D Fourier transform and inverse transform operator are $m^2 \times m^2$ matrices defined as \mathbf{F} and \mathbf{F}^{-1} . $|\cdot|$, $|\cdot|^2$, $\sqrt{\cdot}$, and \cdot/\cdot are element-wise operators, and the $\text{diag}(\cdot)$ operator puts the entries of a vector into the diagonal of a matrix.

The second step is to rewrite the optimization in vector form using the new parameters. First, the forward model (2.1) can be vectorized as

$$\hat{\mathbf{I}}_\ell = |\mathbf{g}_\ell|^2 = |\mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_\ell \mathbf{O}|^2. \quad (2.8)$$

The amplitude-based cost function (2.2) can be vectorized as

$$\min_{\mathbf{O}} f_A(\mathbf{O}) = \min_{\mathbf{O}} \sum_{\ell} (\sqrt{\mathbf{I}_\ell} - |\mathbf{g}_\ell|)^\dagger (\sqrt{\mathbf{I}_\ell} - |\mathbf{g}_\ell|), \quad (2.9)$$

where the hyperscript \dagger denotes a Hermitian conjugate.

Likewise, the intensity-based cost function (2.5) can be vectorized as

$$\min_{\mathbf{O}} f_I(\mathbf{O}) = \min_{\mathbf{O}} \sum_{\ell} (\mathbf{I}_\ell - |\mathbf{g}_\ell|^2)^\dagger (\mathbf{I}_\ell - |\mathbf{g}_\ell|^2). \quad (2.10)$$

The Poisson likelihood cost function is more complicated to be expressed in vector form. First, we rewrite $|\mathbf{g}_\ell|^2$ as

$$|\mathbf{g}_\ell|^2 = \text{diag}(\bar{\mathbf{g}}_\ell) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_\ell \mathbf{O} = \mathbf{A}_\ell \mathbf{O} = \begin{bmatrix} \mathbf{a}_{\ell,1}^\dagger \\ \vdots \\ \mathbf{a}_{\ell,m^2}^\dagger \end{bmatrix} \mathbf{O}, \quad (2.11)$$

where $\mathbf{A}_\ell = \text{diag}(\bar{\mathbf{g}}_\ell) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_\ell$ is a $m^2 \times n^2$ matrix with $m^2 \times 1$ row vectors, $\mathbf{a}_{\ell,j}^\dagger$, $j = 1, \dots, m^2$, and $\bar{\mathbf{g}}_\ell$ denotes the complex conjugate of vector \mathbf{g}_ℓ . Then the likelihood function can be rewritten as

$$\min_{\mathbf{O}} \mathcal{L}_{\text{Poisson}}(\mathbf{O}) = \sum_{\ell} \sum_j [-I_{\ell,j} \log(\mathbf{a}_{\ell,j}^\dagger \mathbf{O}) + \mathbf{a}_{\ell,j}^\dagger \mathbf{O} + \log(I_{\ell,j}!)]. \quad (2.12)$$

To minimize (2.9), (2.10) or (2.12) using an iterative optimization algorithm, the gradients (and possibly Hessians) of the cost functions need to be calculated, both of which are shown in the next section. Since (2.9), (2.10) and (2.12) are all real-valued functions of a complex vector \mathbf{O} , that means that \mathbf{O} and $\bar{\mathbf{O}}$ should be treated independently in the derivative calculation, which is based on the CR-calculus discussed in [103] and the similar formulation for traditional ptychography discussed in [92].

2.3 Algorithms for Fourier ptychographic microscopy

2.3.1 Solving the phase retrieval optimization problem

To solve an optimization problem, the general flow is summarized in Fig. 2.5. The key component of every algorithm is to calculate the gradient and Hessian of the cost function and form the descent direction. With the descent direction, we can then update the complex field with proper step sizes to get the final reconstruction. Here, we calculate the gradient and Hessian of the amplitude-based, intensity-based, and the Poisson-likelihood-based cost functions for later explanation of various algorithms.

2.3.1.1 Gradient of cost functions

Consider that equations (2.9) and (2.10) can be expressed as

$$\begin{aligned} f_A(\mathbf{O}) &= \sum_{\ell} \mathbf{f}_{A\ell}^{\dagger} \mathbf{f}_{A\ell} \\ f_I(\mathbf{O}) &= \sum_{\ell} \mathbf{f}_{I\ell}^{\dagger} \mathbf{f}_{I\ell}, \end{aligned} \quad (2.13)$$

where $\mathbf{f}_{A\ell} \equiv \sqrt{\mathbf{I}_{\ell}} - |\mathbf{g}_{\ell}|$, and $\mathbf{f}_{I\ell} \equiv \mathbf{I}_{\ell} - |\mathbf{g}_{\ell}|^2$.

Then, calculate the derivative of f_A with respect to \mathbf{O} , and it can then be expressed as

$$\nabla_{\mathbf{O}} f_A(\mathbf{O}) = \sum_{\ell} \left[\frac{\partial(\mathbf{f}_{A\ell}^{\dagger} \mathbf{f}_{A\ell})}{\partial \mathbf{O}} \right]^{\dagger} = \sum_{\ell} \left[\frac{\partial(\mathbf{f}_{A\ell}^{\dagger} \mathbf{f}_{A\ell})}{\partial \mathbf{f}_{A\ell}} \frac{\partial \mathbf{f}_{A\ell}}{\partial \mathbf{O}} + \frac{\partial(\mathbf{f}_{A\ell}^{\dagger} \mathbf{f}_{A\ell})}{\partial \mathbf{f}_{A\ell}^{\dagger}} \frac{\partial \mathbf{f}_{A\ell}^{\dagger}}{\partial \mathbf{O}} \right]^{\dagger}. \quad (2.14)$$

Using $|\mathbf{g}_{\ell}|^2 = \text{diag}(\bar{\mathbf{g}}_{\ell}) \mathbf{g}_{\ell}$ and $|\mathbf{g}_{\ell}| = (|\mathbf{g}_{\ell}|^2)^{1/2}$, two chain rule parts in (2.14) are calculated as

$$\begin{aligned} \frac{\partial(\mathbf{f}_{A\ell}^{\dagger} \mathbf{f}_{A\ell})}{\partial \mathbf{f}_{A\ell}} &= \mathbf{f}_{A\ell}^{\dagger} = \mathbf{f}_{A\ell} \\ \frac{\partial(\mathbf{f}_{A\ell}^{\dagger} \mathbf{f}_{A\ell})}{\partial \mathbf{f}_{A\ell}^{\dagger}} &= \mathbf{f}_{A\ell} = \mathbf{f}_{A\ell}^{\dagger} \\ \frac{\partial \mathbf{f}_{A\ell}}{\partial \mathbf{O}} &= \frac{\partial \mathbf{f}_{A\ell}^{\dagger}}{\partial \mathbf{O}} = -\frac{\partial(|\mathbf{g}_{\ell}|^2)^{1/2}}{\partial(|\mathbf{g}_{\ell}|^2)} \frac{\partial(\text{diag}(\bar{\mathbf{g}}_{\ell}) \mathbf{g}_{\ell})}{\partial \mathbf{O}} = -\frac{1}{2} \text{diag} \left(\frac{\bar{\mathbf{g}}_{\ell}}{|\mathbf{g}_{\ell}|} \right) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell}, \end{aligned} \quad (2.15)$$

if \mathbf{g}_{ℓ} does not contain any zero entries for $\ell = 1, \dots, N_{img}$.

By plugging these three terms into (2.14), the gradient of f_A with respect to \mathbf{O} becomes

$$\begin{aligned} \nabla_{\mathbf{O}} f_A(\mathbf{O}) &= -\sum_{\ell} \mathbf{Q}_{\ell}^{\dagger} \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \text{diag} \left(\frac{\mathbf{g}_{\ell}}{|\mathbf{g}_{\ell}|} \right) (\sqrt{\mathbf{I}_{\ell}} - |\mathbf{g}_{\ell}|) \\ &= -\sum_{\ell} \mathbf{Q}_{\ell}^{\dagger} \text{diag}(\bar{\mathbf{P}}) \left(\mathbf{F} \text{diag} \left(\frac{\sqrt{\mathbf{I}_{\ell}}}{|\mathbf{g}_{\ell}|} \right) \mathbf{g}_{\ell} - \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell} \mathbf{O} \right). \end{aligned} \quad (2.16)$$

The gradient for f_I can be calculated in the similar way, and the chain rule part of $\mathbf{f}_{I\ell}$ can be calculated as

$$\frac{\partial \mathbf{f}_{I\ell}}{\partial \mathbf{O}} = \frac{\partial \mathbf{f}_{I\ell}^{\dagger}}{\partial \mathbf{O}} = -\frac{\partial(\text{diag}(\bar{\mathbf{g}}_{\ell}) \mathbf{g}_{\ell})}{\partial \mathbf{O}} = -\text{diag}(\bar{\mathbf{g}}_{\ell}) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell}. \quad (2.17)$$

With (2.17), it is clear to express the gradient of f_I as

$$\nabla_{\mathbf{O}} f_I(\mathbf{O}) = \sum_{\ell} \left[\frac{\partial(\mathbf{f}_{I\ell}^{\dagger} \mathbf{f}_{I\ell})}{\partial \mathbf{f}_{I\ell}} \frac{\partial \mathbf{f}_{I\ell}}{\partial \mathbf{O}} + \frac{\partial(\mathbf{f}_{I\ell}^{\dagger} \mathbf{f}_{I\ell})}{\partial \mathbf{f}_{I\ell}^{\dagger}} \frac{\partial \mathbf{f}_{I\ell}^{\dagger}}{\partial \mathbf{O}} \right]^{\dagger} = -2 \sum_{\ell} \mathbf{Q}_{\ell}^{\dagger} \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \text{diag}(\mathbf{g}_{\ell}) (\mathbf{I}_{\ell} - |\mathbf{g}_{\ell}|^2). \quad (2.18)$$

The calculation of gradient of $\mathcal{L}_{\text{Poisson}}(\mathbf{O})$ with respect to \mathbf{O} is different from the other two. With the expression (2.12), the gradient of Poisson likelihood function can be calculated as

$$\begin{aligned}
 \nabla_{\mathbf{O}} \mathcal{L}_{\text{Poisson}}(\mathbf{O}) &= \left(\frac{\partial \mathcal{L}_{\text{Poisson}}}{\partial \mathbf{O}} \right)^\dagger = \left(\sum_{\ell} \sum_j \left[-\frac{I_{\ell,j}}{\mathbf{a}_{\ell,j}^\dagger \mathbf{O}} \mathbf{a}_{\ell,j}^\dagger + \mathbf{a}_{\ell,j}^\dagger \right] \right)^\dagger \\
 &= - \left(\sum_{\ell} \sum_j \left[I_{\ell,j} - \mathbf{a}_{\ell,j}^\dagger \mathbf{O} \right] \frac{1}{\mathbf{a}_{\ell,j}^\dagger \mathbf{O}} \mathbf{a}_{\ell,j}^\dagger \right)^\dagger \\
 &= - \left(\sum_{\ell} (\mathbf{I}_{\ell} - |\mathbf{g}_{\ell}|^2)^\dagger \text{diag} \left(\frac{1}{|\mathbf{g}_{\ell}|^2} \right) \text{diag}(\bar{\mathbf{g}}_{\ell}) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell} \right)^\dagger \\
 &= - \sum_{\ell} \mathbf{Q}_{\ell}^\dagger \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \text{diag} \left(\frac{\mathbf{g}_{\ell}}{|\mathbf{g}_{\ell}|^2} \right) (\mathbf{I}_{\ell} - |\mathbf{g}_{\ell}|^2). \tag{2.19}
 \end{aligned}$$

This is equivalent to the gradient of the *intensity-based* cost function with added weight $1/|\mathbf{g}_{\ell}|^2$ to the component from each image. In addition, this gradient is very similar to that from the *amplitude-based* cost function in (2.16).

Since we have gradients for all cost functions, the updating equation for the gradient descent method can then be expressed as

$$\mathbf{O}^{(i+1)} = \mathbf{O}^{(i)} - \alpha^{(i)} \nabla_{\mathbf{O}} f(\mathbf{O}^{(i)}), \tag{2.20}$$

where i denotes the iteration number, α is the step size chosen by the line search algorithm, and $f(\mathbf{O})$ can be either *intensity-based* or *amplitude-based* cost function.

Looking at $\nabla_{\mathbf{O}} f_A(\mathbf{O})$, $\nabla_{\mathbf{O}} f_I(\mathbf{O})$ and $\nabla_{\mathbf{O}} \mathcal{L}_{\text{Poisson}}(\mathbf{O})$, they all contain the term $\mathbf{Q}_{\ell}^\dagger \text{diag}(\bar{\mathbf{P}})$ followed by a residual term. The residual term basically finds the difference between the estimate and the measurement. This difference carries the information to update the previous estimation. Since each measurement carries the information for a specific region in the Fourier space, the $\mathbf{Q}_{\ell}^\dagger \text{diag}(\bar{\mathbf{P}})$ term brings this updated information back to the right place corresponding to some spatial frequency. For $\nabla_{\mathbf{O}} f_A(\mathbf{O})$, the first term in the residual shows the replacement of the amplitude in the real domain, which is the projection from the estimate to the modulus space. Thus, the gradient descent method using the *amplitude-based* cost function is similar to the projection-based phase retrieval solver.

2.3.1.2 Hessian of cost functions

The second-order Taylor expansion on an arbitrary real function $f(\mathbf{c})$ with a complex vector $\mathbf{c} = (\mathbf{O}^T, \bar{\mathbf{O}}^T)^T$ at certain point $\mathbf{c}_0 = (\mathbf{O}_0^T, \bar{\mathbf{O}}_0^T)^T$ can be written as [103]

$$f(\mathbf{c}) \approx f(\mathbf{c}_0) + \nabla f(\mathbf{c}_0)^\dagger (\mathbf{c} - \mathbf{c}_0) + \frac{1}{2} (\mathbf{c} - \mathbf{c}_0)^\dagger \mathbf{H}_{\mathbf{c}\mathbf{c}}(\mathbf{c}_0) (\mathbf{c} - \mathbf{c}_0), \tag{2.21}$$

where the matrix $\mathbf{H}_{\mathbf{cc}}$ is the Hessian of $f(\mathbf{c})$. For the case of a single-value function, the second-order term in the Taylor expansion denotes the curvature of the function at that expansion point. Thus, this Hessian matrix similarly contains the curvature information of the original multi-variate function.

If the Hessian is a diagonal matrix, each diagonal entry denotes the curvature in each corresponding dimension. If the Hessian is not diagonal, a coordinate transformation can be found to make the Hessian diagonal by using eigenvalue decomposition. For a convex problem, the Hessian is positive semidefinite. The curvatures of the cost function in different dimensions are always nonnegative. A standard optimization process can lead to a global minimum. However, if the problem is non-convex, a standard optimization process will probably lead to a local minimum. Calculating the Hessian of a cost function is useful either to examine the optimization process or to speed up the convergence rate by using Newton's method.

From [92, 103], the definition for the Hessian of a real-value function with multiple complex variables is a $2n^2 \times 2n^2$ matrix and can be expressed as

$$\mathbf{H}_{\mathbf{cc}} = \begin{bmatrix} \mathbf{H}_{\mathbf{oo}} & \mathbf{H}_{\bar{\mathbf{o}}\mathbf{o}} \\ \mathbf{H}_{\mathbf{o}\bar{\mathbf{o}}} & \mathbf{H}_{\bar{\mathbf{o}}\bar{\mathbf{o}}} \end{bmatrix}, \quad (2.22)$$

where each component $n^2 \times n^2$ matrices can be further calculated as

$$\begin{aligned} \mathbf{H}_{\mathbf{oo}} &= \frac{\partial}{\partial \mathbf{O}} \left(\frac{\partial f}{\partial \mathbf{O}} \right)^\dagger, \quad \mathbf{H}_{\bar{\mathbf{o}}\mathbf{o}} = \frac{\partial}{\partial \bar{\mathbf{O}}} \left(\frac{\partial f}{\partial \mathbf{O}} \right)^\dagger \\ \mathbf{H}_{\mathbf{o}\bar{\mathbf{o}}} &= \frac{\partial}{\partial \mathbf{O}} \left(\frac{\partial f}{\partial \bar{\mathbf{O}}} \right)^\dagger, \quad \mathbf{H}_{\bar{\mathbf{o}}\bar{\mathbf{o}}} = \frac{\partial}{\partial \bar{\mathbf{O}}} \left(\frac{\partial f}{\partial \bar{\mathbf{O}}} \right)^\dagger. \end{aligned} \quad (2.23)$$

Similar to the calculation of the gradient, the components of the Hessians for *amplitude-based*, *intensity-based*, and *Poisson-likelihood-based* cost functions can be calculated by taking an additional derivative on the gradient of the cost functions. The components of the Hessian for the amplitude-based cost function are

$$\begin{aligned} \mathbf{H}_{\mathbf{oo}}^A &= \sum_{\ell} \mathbf{Q}_{\ell}^\dagger \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \left[\mathbf{1} - \frac{1}{2} \text{diag} \left(\frac{\sqrt{\mathbf{I}_{\ell}}}{|\mathbf{g}_{\ell}|} \right) \right] \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell} \\ \mathbf{H}_{\bar{\mathbf{o}}\mathbf{o}}^A &= \frac{1}{2} \sum_{\ell} \mathbf{Q}_{\ell}^\dagger \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \text{diag} \left(\frac{\sqrt{\mathbf{I}_{\ell}} \mathbf{g}_{\ell}^2}{|\mathbf{g}_{\ell}|^3} \right) \bar{\mathbf{F}}^{-1} \text{diag}(\bar{\mathbf{P}}) \bar{\mathbf{Q}}_{\ell} \\ \mathbf{H}_{\mathbf{o}\bar{\mathbf{o}}}^A &= \frac{1}{2} \sum_{\ell} \mathbf{Q}_{\ell}^T \text{diag}(\mathbf{P}) \bar{\mathbf{F}} \text{diag} \left(\frac{\sqrt{\mathbf{I}_{\ell}} \bar{\mathbf{g}}_{\ell}^2}{|\mathbf{g}_{\ell}|^3} \right) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell} \\ \mathbf{H}_{\bar{\mathbf{o}}\bar{\mathbf{o}}}^A &= \sum_{\ell} \mathbf{Q}_{\ell}^T \text{diag}(\mathbf{P}) \bar{\mathbf{F}} \left[\mathbf{1} - \frac{1}{2} \text{diag} \left(\frac{\sqrt{\mathbf{I}_{\ell}}}{|\mathbf{g}_{\ell}|} \right) \right] \bar{\mathbf{F}}^{-1} \text{diag}(\bar{\mathbf{P}}) \bar{\mathbf{Q}}_{\ell}, \end{aligned} \quad (2.24)$$

where $\mathbf{1}$ is the $m^2 \times m^2$ identity matrix.

Likewise, the Hessian of the intensity-based cost function is

$$\begin{aligned}
 \mathbf{H}_{\mathbf{O}\mathbf{O}}^I &= 2 \sum_{\ell} \mathbf{Q}_{\ell}^{\dagger} \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \text{diag}(2|\mathbf{g}_{\ell}|^2 - \mathbf{I}_{\ell}) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell} \\
 \mathbf{H}_{\bar{\mathbf{O}}\bar{\mathbf{O}}}^I &= 2 \sum_{\ell} \mathbf{Q}_{\ell}^{\dagger} \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \text{diag}(\mathbf{g}_{\ell}^2) \bar{\mathbf{F}}^{-1} \text{diag}(\bar{\mathbf{P}}) \bar{\mathbf{Q}}_{\ell} \\
 \mathbf{H}_{\mathbf{O}\bar{\mathbf{O}}}^I &= 2 \sum_{\ell} \mathbf{Q}_{\ell}^T \text{diag}(\mathbf{P}) \bar{\mathbf{F}} \text{diag}(\bar{\mathbf{g}}_{\ell}^2) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell} \\
 \mathbf{H}_{\bar{\mathbf{O}}\mathbf{O}}^I &= 2 \sum_{\ell} \mathbf{Q}_{\ell}^T \text{diag}(\mathbf{P}) \bar{\mathbf{F}} \text{diag}(2|\mathbf{g}_{\ell}|^2 - \mathbf{I}_{\ell}) \bar{\mathbf{F}}^{-1} \text{diag}(\bar{\mathbf{P}}) \bar{\mathbf{Q}}_{\ell}.
 \end{aligned} \tag{2.25}$$

Finally, the Hessian of the Poisson likelihood cost function is

$$\begin{aligned}
 \mathbf{H}_{\mathbf{O}\mathbf{O}}^P &= \sum_{\ell} \mathbf{Q}_{\ell}^{\dagger} \text{diag}(|\mathbf{P}|^2) \mathbf{Q}_{\ell} \\
 \mathbf{H}_{\bar{\mathbf{O}}\bar{\mathbf{O}}}^P &= \sum_{\ell} \mathbf{Q}_{\ell}^{\dagger} \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \text{diag}\left(\frac{\mathbf{I}_{\ell} \mathbf{g}_{\ell}^2}{|\mathbf{g}_{\ell}|^4}\right) \bar{\mathbf{F}}^{-1} \text{diag}(\bar{\mathbf{P}}) \bar{\mathbf{Q}}_{\ell} \\
 \mathbf{H}_{\mathbf{O}\bar{\mathbf{O}}}^P &= \sum_{\ell} \mathbf{Q}_{\ell}^T \text{diag}(\mathbf{P}) \bar{\mathbf{F}} \text{diag}\left(\frac{\mathbf{I}_{\ell} \bar{\mathbf{g}}_{\ell}^2}{|\mathbf{g}_{\ell}|^4}\right) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell} \\
 \mathbf{H}_{\bar{\mathbf{O}}\mathbf{O}}^P &= \sum_{\ell} \mathbf{Q}_{\ell}^T \text{diag}(|\mathbf{P}|^2) \bar{\mathbf{Q}}_{\ell}.
 \end{aligned} \tag{2.26}$$

In general, Newton's method, which is the second-order method using the inversion of Hessian matrix, is preferred in solving nonlinear least square problems because of its fast convergence and stability compared to the first-order methods such as gradient descent. The updating equation for Newton's method can be expressed as

$$\begin{bmatrix} \mathbf{O}^{(i+1)} \\ \bar{\mathbf{O}}^{(i+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{O}^{(i)} \\ \bar{\mathbf{O}}^{(i)} \end{bmatrix} - \alpha^{(i)} \mathbf{H}_{\text{cc}}^{-1} \begin{bmatrix} \nabla_{\mathbf{O}} f(\mathbf{O}^{(i)}) \\ \nabla_{\bar{\mathbf{O}}} f(\bar{\mathbf{O}}^{(i)}) \end{bmatrix}. \tag{2.27}$$

2.3.2 First-order methods

2.3.2.1 Sequential gradient descent method (GS algorithm) [20, 93]

For the implementation in [20, 93], the algorithm aims to optimize the amplitude-based cost function (2.9). It is the simplest to implement and, in this case, equivalent to the Gerchberg-Saxton approach of simply replacing known information in real and Fourier space. Since the sequential strategy treats a single image as an optimization problem, the cost function for each problem is just one component of Eq. (2.9) and is defined as

$$f_{A,\ell}(\mathbf{O}) = (\sqrt{\mathbf{I}_{\ell}} - |\mathbf{g}_{\ell}|)^{\dagger} (\sqrt{\mathbf{I}_{\ell}} - |\mathbf{g}_{\ell}|), \tag{2.28}$$

where ℓ denotes the index of each measurement.

The derivative of this cost function is thus a component of Eq. (2.16) and can be expressed as

$$\nabla_{\mathbf{O}} f_{A,\ell}(\mathbf{O}) = -\mathbf{Q}_\ell^\dagger \text{diag}(\bar{\mathbf{P}}) \left[\mathbf{F} \text{diag} \left(\frac{\sqrt{\mathbf{I}_\ell}}{|\mathbf{g}_\ell|} \right) \mathbf{g}_\ell - \text{diag}(\mathbf{P}) \mathbf{Q}_\ell \mathbf{O} \right]. \quad (2.29)$$

The update equation for this sequential amplitude-based algorithm is then a gradient descent with the descent direction given by Eq. (2.29) and step size $1/|\mathbf{P}|_{\max}^2$:

$$\mathbf{O}^{(i,\ell+1)} = \mathbf{O}^{(i,\ell)} - \frac{1}{|\mathbf{P}|_{\max}^2} \nabla_{\mathbf{O}} f_{A,\ell+1}(\mathbf{O}^{(i,\ell)}), \quad (2.30)$$

where i indicates the iteration number, which goes to $i + 1$ after running through all the measurements from $\ell = 1$ to $\ell = N_{img}$. This algorithm adopts the alternating projection phase retrieval approach. The first projection in the real domain is the amplitude replacement operation $\text{diag} \left(\frac{\sqrt{\mathbf{I}_\ell}}{|\mathbf{g}_\ell|} \right) \mathbf{g}_\ell$, and the second projection is to project the previous estimated Fourier region $\text{diag}(\mathbf{P}) \mathbf{Q}_\ell \mathbf{O}$ onto the updated Fourier region $\mathbf{F} \text{diag} \left(\frac{\sqrt{\mathbf{I}_\ell}}{|\mathbf{g}_\ell|} \right) \mathbf{g}_\ell$.

It is worth noting that the algorithm in [20] directly replaces $\mathbf{F} \text{diag} \left(\frac{\sqrt{\mathbf{I}_\ell}}{|\mathbf{g}_\ell|} \right) \mathbf{g}_\ell$ in the Fourier domain at each sub-iteration. A similar algorithm in [93], introduced for simultaneous aberration recovery, has the same form as Eq. (2.30) that implements gradient descent in the Fourier domain. However, when there is no pupil estimation, then \mathbf{P} becomes a pure support function with one inside the support and zero outside. In this situation, these two algorithms become exactly the same, and thus we refer to both as sequential gradient descent or Gerchberg-Saxton algorithm.

2.3.2.2 Wirtinger-flow algorithm [94, 95]

The Wirtinger flow optimization framework was originally proposed to iteratively solve the coded-mask phase retrieval problem using nonlinear optimization [94]. It is a gradient descent method implemented with a special initialization and special step sizes. For the FPM implementation described in [95], the *intensity-based* cost function is used. Thus, the update equation for the object transmission function \mathbf{O} can be expressed as

$$\mathbf{O}^{(i+1)} = \mathbf{O}^{(i)} - \alpha^{(i)} \nabla_{\mathbf{O}} f_I(\mathbf{O}^{(i)}), \quad (2.31)$$

where the step size is calculated by

$$\alpha^{(i)} = \frac{\min(1 - e^{-i/i_0}, \theta_{\max})}{(\mathbf{O}^{(0)})^\dagger \mathbf{O}^{(0)}}, \quad (2.32)$$

where $\nabla_{\mathbf{O}} f_I(\mathbf{O}^{(i)})$ is the gradient of the intensity-based cost function calculated in (2.18), and i_0 and θ_{\max} are user-chosen parameters to calculate the step size.

In the previously proposed FPM implementation of Wirtinger flow [95], the algorithm deviates somewhat from the original theory proposed in [94]. First, there is an additional term in the cost function to deal with additive noise. Second, the initialization used in [95] is not the proposed one in [94], but rather a low-resolution captured image. So the algorithm in [95] is essentially a gradient descent method with the special step size based on the intensity-based cost function and is not guaranteed to converge to the global minimum.

The Wirtinger flow algorithm can be implemented with different cost functions simply by replacing the original *intensity-based* gradient with the other gradients derived in the Appendix. For comparison, we have implemented the Wirtinger flow algorithm using all three of the cost functions described here: *amplitude-based*, *intensity-based* and *Poisson-likelihood-based*. The results are compared in Fig. 2.9 with experimental data and Section 2.4 with simulated data.

2.3.3 Second-order methods

Beyond first-order, a second-order optimization method can improve the convergence speed and stability of the algorithm, especially for nonlinear and non-convex problems. Second-order methods (e.g. Newton's method) use both the first and second derivatives (Hessian) of the cost function to create a better update at each iteration. As a result, they generally require fewer iterations and move more directly towards the solution. The difficulty of second-order implementations is in computing the Hessian matrix, whose size scales quadratically with the size of the image. As a result, approximations to the Hessian are often used (known as *quasi-Newton* methods) to trade performance for computational efficiency.

2.3.3.1 Sequential Gauss-Newton method [21]

First, we look at a Gauss-Newton method based on the *amplitude-based* cost function, which approximates the Hessian matrix as a multiplication of its Jacobian matrix:

$$\begin{aligned} \mathbf{H}_{\mathbf{cc},\ell}^A &\approx \left(\frac{\partial \mathbf{f}_{A\ell}}{\partial \mathbf{c}} \right)^\dagger \left(\frac{\partial \mathbf{f}_{A\ell}}{\partial \mathbf{c}} \right) \\ &= \begin{bmatrix} \frac{1}{2} \mathbf{Q}_\ell^\dagger \text{diag}(|\mathbf{P}|^2) \mathbf{Q}_\ell & \mathbf{Q}_\ell^\dagger \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \text{diag} \left(\frac{\mathbf{g}_\ell^2}{|\mathbf{g}_\ell|^2} \right) \bar{\mathbf{F}}^{-1} \text{diag}(\bar{\mathbf{P}}) \bar{\mathbf{Q}}_\ell \\ \mathbf{Q}_\ell^T \text{diag}(\mathbf{P}) \bar{\mathbf{F}} \text{diag} \left(\frac{\bar{\mathbf{g}}_\ell^2}{|\bar{\mathbf{g}}_\ell|^2} \right) \mathbf{F}^{-1} \text{diag}(\mathbf{P}) \mathbf{Q}_\ell & \frac{1}{2} \mathbf{Q}_\ell^T \text{diag}(|\mathbf{P}|^2) \bar{\mathbf{Q}}_\ell \end{bmatrix}, \end{aligned} \quad (2.33)$$

where $\mathbf{c} = (\mathbf{O}^T, \bar{\mathbf{O}}^T)^T$ (See Appendix). Since the inversion of this Hessian matrix requires very high computational cost, we approximate the Hessian by dropping all the off-diagonal terms of the Hessian matrix. Further, the inversion of the Hessian matrix may be an ill-posed problem, so a constant regularizer is adopted. In the end, the approximated Hessian

inversion becomes

$$(\mathbf{H}_{\mathbf{cc},\ell}^A)^{-1} \approx \begin{bmatrix} 2\mathbf{Q}_\ell^\dagger \text{diag}\left(\frac{1}{|\mathbf{P}|^2+\Delta}\right) \mathbf{Q}_\ell & \mathbf{0} \\ \mathbf{0} & 2\mathbf{Q}_\ell^T \text{diag}\left(\frac{1}{|\mathbf{P}|^2+\Delta}\right) \bar{\mathbf{Q}}_\ell \end{bmatrix}, \quad (2.34)$$

where Δ is a constant vector with all the entries equal to a constant regularizer δ over all pixels.

By applying Newton's update, Eq. (2.27), with this approximated Hessian inversion, the new estimate of \mathbf{O} can be expressed as

$$\begin{bmatrix} \mathbf{O}^{(i,\ell+1)} \\ \bar{\mathbf{O}}^{(i,\ell+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{O}^{(i,\ell)} \\ \bar{\mathbf{O}}^{(i,\ell)} \end{bmatrix} - \begin{bmatrix} \mathbf{Q}_\ell^\dagger \text{diag}\left(\frac{|\mathbf{P}|}{|\mathbf{P}|_{\max}}\right) \mathbf{Q}_\ell & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_\ell^T \text{diag}\left(\frac{|\mathbf{P}|}{|\mathbf{P}|_{\max}}\right) \bar{\mathbf{Q}}_\ell \end{bmatrix} (\mathbf{H}_{\mathbf{cc},\ell}^A)^{-1} \begin{bmatrix} \nabla_{\mathbf{O}} f_{A,\ell+1}(\mathbf{O}^{(i,\ell)}) \\ \nabla_{\bar{\mathbf{O}}} f_{A,\ell+1}(\mathbf{O}^{(i,\ell)}) \end{bmatrix} \quad (2.35)$$

where the $\text{diag}\left(\frac{|\mathbf{P}|}{|\mathbf{P}|_{\max}}\right)$ part is the step size for this descent direction. Note that when \mathbf{P} is a constant having either 0 or 1 values, this method is reduced to the sequential gradient descent method with a tunable regularizer δ . In practice, however, we also simultaneously update \mathbf{P} (see Section 2.5.1), so the second-order optimization procedure becomes more crucial.

2.3.3.2 Global Newton's method

Since we expect second-order methods to perform better than first-order, and also global methods to be more stable than sequential, we propose a new global second-order (Newton's) method, and show the results compared against other methods. For completeness, we implement all three of the amplitude, intensity, and Poisson-likelihood-based cost functions, showing that the amplitude and Poisson-likelihood-based cost functions indeed perform better. The difficult step in deriving a Newton's method for this problem is in calculating the gradients and Hessians of the cost functions directly, without approximations. In the Appendix, we show our derivation, and in this section we use the results with a typical Newton's update equation:

$$\begin{bmatrix} \mathbf{O}^{(i+1)} \\ \bar{\mathbf{O}}^{(i+1)} \end{bmatrix} = \begin{bmatrix} \mathbf{O}^{(i)} \\ \bar{\mathbf{O}}^{(i)} \end{bmatrix} - \alpha^{(i)} (\mathbf{H}_{\mathbf{cc}})^{-1} \begin{bmatrix} \nabla_{\mathbf{O}} f(\mathbf{O}^{(i)}) \\ \nabla_{\bar{\mathbf{O}}} f(\mathbf{O}^{(i)}) \end{bmatrix}. \quad (2.36)$$

The inverse of the Hessian matrix, $(\mathbf{H}_{\mathbf{cc}})^{-1}$, is solved efficiently by a conjugate gradient matrix inversion iterative solver as described in [104]. $\alpha^{(i)}$ is determined by the backtracking line search algorithm at each iteration, as described in [96]. The exact form of the cost function and the Hessian depends on the algorithm used. For *amplitude-based* Newton's algorithm, $f(\mathbf{O}) = f_A(\mathbf{O})$ and $\mathbf{H}_{\mathbf{cc}} = \mathbf{H}_{\mathbf{cc}}^A$; for *intensity-based* Newton's algorithm, $f(\mathbf{O}) = f_I(\mathbf{O})$ and $\mathbf{H}_{\mathbf{cc}} = \mathbf{H}_{\mathbf{cc}}^I$; for *Poisson-likelihood-based* Newton's algorithm, $f(\mathbf{O}) = \mathcal{L}_{\text{Gaussian}}(\mathbf{O})$ and $\mathbf{H}_{\mathbf{cc}} = \mathbf{H}_{\mathbf{cc}}^P$.

Since this problem is nonlinear, its Hessian can be negative definite for some iterations. For this case, the conjugate gradient solver can not directly be applied to solve this inverse problem and the gradient descent direction is used instead to prevent this situation in that specific iteration. In [92], they prove the Hessian of the intensity-based and amplitude-based cost functions for ptychography is usually positive-semi-definite if the initial guess is close to the solution. The same proof can be applied to the Fourier ptychography case. This is why in the real implementation the negative definiteness of the Hessian matrix is seldom encountered.

2.3.4 Convex-based method

2.3.4.1 PhaseLift algorithm [97–101]

The PhaseLift formulation for phase retrieval is conceptually quite different than the previous methods described here. The idea is to lift the non-convex problem into a higher-dimensional space in which it is convex, thereby guaranteeing convergence to the global solution. To do this, the cost function of \mathbf{O} is reformulated into that of a rank-1 matrix $\mathbf{X} = \mathbf{O}\mathbf{O}^\dagger$ and the goal is to estimate \mathbf{X} instead of \mathbf{O} . The process of reformulation can be expressed as [101]

$$\begin{aligned} \mathbf{g} &= \begin{bmatrix} \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{N_{img}} \end{bmatrix} = \begin{bmatrix} \mathbf{F}^{-1} & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \mathbf{F}^{-1} \end{bmatrix} \begin{bmatrix} \text{diag}(\mathbf{P}) & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \text{diag}(\mathbf{P}) \end{bmatrix} \begin{bmatrix} \mathbf{Q}_1 \\ \vdots \\ \mathbf{Q}_{N_{img}} \end{bmatrix} \mathbf{O} \\ &= \mathbf{D}\mathbf{O} = \begin{bmatrix} \mathbf{d}_1^\dagger \\ \vdots \\ \mathbf{d}_{N_{img}m^2}^\dagger \end{bmatrix} \mathbf{O}, \end{aligned} \quad (2.37)$$

where \mathbf{D} is an $N_{img}m^2 \times N_{img}m^2$ operator combining the inverse Fourier transform, pupil cropping, and the downsampling operation with row vectors denoted by \mathbf{d}_j^\dagger .

Hence, the estimated intensity $|\mathbf{g}|^2$ as a function of \mathbf{X} can be expressed

$$|\mathbf{g}|^2 = \begin{bmatrix} \mathbf{O}^\dagger \mathbf{d}_1 \mathbf{d}_1^\dagger \mathbf{O} \\ \vdots \\ \mathbf{O}^\dagger \mathbf{d}_{N_{img}m^2} \mathbf{d}_{N_{img}m^2}^\dagger \mathbf{O} \end{bmatrix} = \begin{bmatrix} \text{Tr}(\mathbf{d}_1 \mathbf{d}_1^\dagger \mathbf{O}\mathbf{O}^\dagger) \\ \vdots \\ \text{Tr}(\mathbf{d}_{N_{img}m^2} \mathbf{d}_{N_{img}m^2}^\dagger \mathbf{O}\mathbf{O}^\dagger) \end{bmatrix} = \begin{bmatrix} \text{Tr}(\mathbf{D}_1 \mathbf{X}) \\ \vdots \\ \text{Tr}(\mathbf{D}_{N_{img}m^2} \mathbf{X}) \end{bmatrix} = \mathcal{A}(\mathbf{X}), \quad (2.38)$$

where \mathcal{A} is a linear operator transforming \mathbf{X} into $|\mathbf{g}|^2$. In Section 2.2.2, we discussed three different cost functions. Only the intensity-based and Poisson-likelihood-based cost functions are convex on the estimated intensity, $\hat{I}_\ell(\mathbf{r})$, which is a component of $\mathcal{A}(\mathbf{X})$. Thus, the intensity-based and Poisson-likelihood-based cost functions can be turned into a convex function on \mathbf{X} through this transformation. For the implementation in [101], by defining $\mathbf{I} = [\mathbf{I}_1^T, \dots, \mathbf{I}_{N_{img}}^T]^T$, the intensity-based cost function can be expressed as

$$\begin{aligned} f_I(\mathbf{X}) &= (\mathbf{I} - |\mathbf{g}|^2)^\dagger (\mathbf{I} - |\mathbf{g}|^2) \\ &= (\mathbf{I} - \mathcal{A}(\mathbf{X}))^\dagger (\mathbf{I} - \mathcal{A}(\mathbf{X})). \end{aligned} \quad (2.39)$$

Since \mathbf{X} is a rank-1 matrix, we then minimize the rank of \mathbf{X} subject to $\mathbf{I} = \mathcal{A}(\mathbf{X})$. However, the rank minimization problem is NP-hard. Therefore, a convex relaxation [97–99] is used instead to transform the problem into a trace minimization problem. Under this relaxation, the optimization problem becomes

$$\min_{\mathbf{X}} f'_I(\mathbf{X}) = \min_{\mathbf{X}} (\mathbf{I} - \mathcal{A}(\mathbf{X}))^\dagger (\mathbf{I} - \mathcal{A}(\mathbf{X})) + \alpha \text{Tr}(\mathbf{X}), \quad (2.40)$$

where α is a regularization variable that depends on the noise level.

The problem with this new approach is that by increasing the dimensionality of the problem, the size of the matrix \mathbf{X} has become $n^2 \times n^2$, which is too large to store and calculate eigenvalue decomposition on a normal computer. To avoid these computational problems, we do not directly solve (2.40), but rather apply a factorization to $\mathbf{X} = \mathbf{R}\mathbf{R}^\dagger$, where \mathbf{R} is an $n^2 \times k$ matrix. \mathbf{X} is a rank-1 matrix so k is set to be 1 (\mathbf{R} becomes \mathbf{O}). This new problem is then solved effectively using the augmented Lagrangian multiplier, by modifying the original cost function [100, 101]

$$\min_{\mathbf{R}} f_{AL,I}(\mathbf{R}) = \min_{\mathbf{R}} \frac{\sigma}{2} (\mathbf{I} - \mathcal{A}(\mathbf{R}\mathbf{R}^\dagger))^\dagger (\mathbf{I} - \mathcal{A}(\mathbf{R}\mathbf{R}^\dagger)) + \mathbf{y}^T (\mathbf{I} - \mathcal{A}(\mathbf{R}\mathbf{R}^\dagger)) + \text{Tr}(\mathbf{R}\mathbf{R}^\dagger), \quad (2.41)$$

where \mathbf{y} , $N_{img}m^2 \times 1$ vector, is the Lagrangian multiplier, and $\sigma \geq 0$ is the augmented Lagrangian multiplier. Both are parameters that can be tuned to give a better reconstruction. By taking the derivative of this cost function with respect to \mathbf{R} and updating \mathbf{R} in each iteration, the optimization problem can then be solved [100]. Unfortunately, after these modifications, the problem becomes non-convex because of the minimization with respect to \mathbf{R} instead of \mathbf{X} , and thus is no longer provable.

In order to provide a more familiar form for comparing the PhaseLift algorithm to the others discussed in this paper, we define $\mathbf{y} = [\mathbf{y}_1^T, \dots, \mathbf{y}_{N_{img}}^T]^T$, where \mathbf{y}_i is $m^2 \times 1$ vector, so that the minimization problem in Eq. (2.41) becomes

$$\min_{\mathbf{O}} f_{AL,I}(\mathbf{O}) = \min_{\mathbf{O}} \frac{\sigma}{2} \left[\sum_{\ell} (\mathbf{I}_{\ell} - |\mathbf{g}_{\ell}|^2 + \frac{2}{\sigma} \mathbf{y}_{\ell})^\dagger (\mathbf{I}_{\ell} - |\mathbf{g}_{\ell}|^2) \right] + \mathbf{O}^\dagger \mathbf{O}. \quad (2.42)$$

Now, we see that the PhaseLift implementation is essentially an intensity-based cost function with an additional constraint that may deal better with noise.

The corresponding derivative of the cost function is calculated as in the previous section:

$$\nabla_{\mathbf{O}} f_{AL,I}(\mathbf{O}) = -\sigma \sum_{\ell} \mathbf{Q}_{\ell}^\dagger \text{diag}(\bar{\mathbf{P}}) \mathbf{F} \text{diag}(\mathbf{g}_{\ell}) \left(\mathbf{I}_{\ell} - |\mathbf{g}_{\ell}|^2 + \frac{1}{\sigma} \mathbf{y}_{\ell} \right) + \mathbf{O}. \quad (2.43)$$

When σ is large compared to the component of \mathbf{y}_{ℓ} and \mathbf{O} , the factorized PhaseLift formulation with rank-1 \mathbf{X} is equivalent to the intensity-based optimization problem discussed in the previous section. To solve this optimization problem, a quasi-Newton algorithm called L-BFGS (Limited-memory Broyden-Fletcher-Goldfarb-Shanno) method [96], which is a second-order method using an approximated Hessian inversion from previous gradients, is adopted.

We note that although the PhaseLift algorithm can also be implemented with the *Poisson-likelihood-based* cost function, the algorithm in the rank-1 case is equivalent to our global Newton’s method discussed in Section 2.3.3.2 for the same reason as in the above analysis.

2.4 Performance analysis of various algorithms

In this section, we compare the algorithms described in Section 2.3 using experimental data, as well as simulated data that mimics the experimental errors described in Section 2.4.1. We find that second-order optimization generally performs better than first-order, while global methods do not give significant improvement over sequential. Further, we explain why the cost function is a key consideration in choosing an algorithm by explaining the cause of the high-frequency artifacts that result from *intensity-based* algorithms. Interestingly, the two model mis-match errors (aberrations and LED mis-alignment) behave similarly to Poisson noise, in that they also give intensity-dependent errors. Hence, the amplitude and Poisson likelihood algorithms are more robust not only to Poisson noise, but also to model mis-match errors.

2.4.1 Possible noise and simulated dataset

Ideally, all algorithms based on the forward model above should give good reconstructions. However, noise and model mis-match errors cause deviations from our forward model. Thus, a noise model that accurately describes the error will be important for noise tolerance. Heuristically, we have identified three experimental non-idealities that cause error: Poisson noise, aberrations and LED mis-alignment. We aim to separate and analyze the artifacts caused by each through controlled simulations that incur only one type of error.

The simulated data (Fig. 2.6) uses the same parameters as our experimental setup, where a 32×32 green LED array (central wavelength $\lambda = 514$ nm) is placed 77 mm above the sample. LEDs are nominally 4 mm apart from each other and only the central 293 LEDs are used, giving a maximum $NA_{illu} = 0.45$. Samples are imaged with a $4\times$ objective lens having $NA_{obj} = 0.1$.

Using our forward model, we simulate four datasets:

1. Ideal data: no noise is added. The object and pupil follow exactly the FPM forward model that is assumed in the algorithm.
2. Poisson noise data: the ideal data is corrupted by Poisson-distributed noise at each pixel. To emphasize the effect and to emulate experiments with lower-performance sensors, we simulate $20\times$ more noise than is present in our experiments (details in Section 2.2.2).
3. Aberrated data: simulated images are corrupted by imaging system aberrations, which are described by the aberrated complex pupil function shown in Fig. 2.6. The pupil function used in these simulations was obtained from experimental measurements.

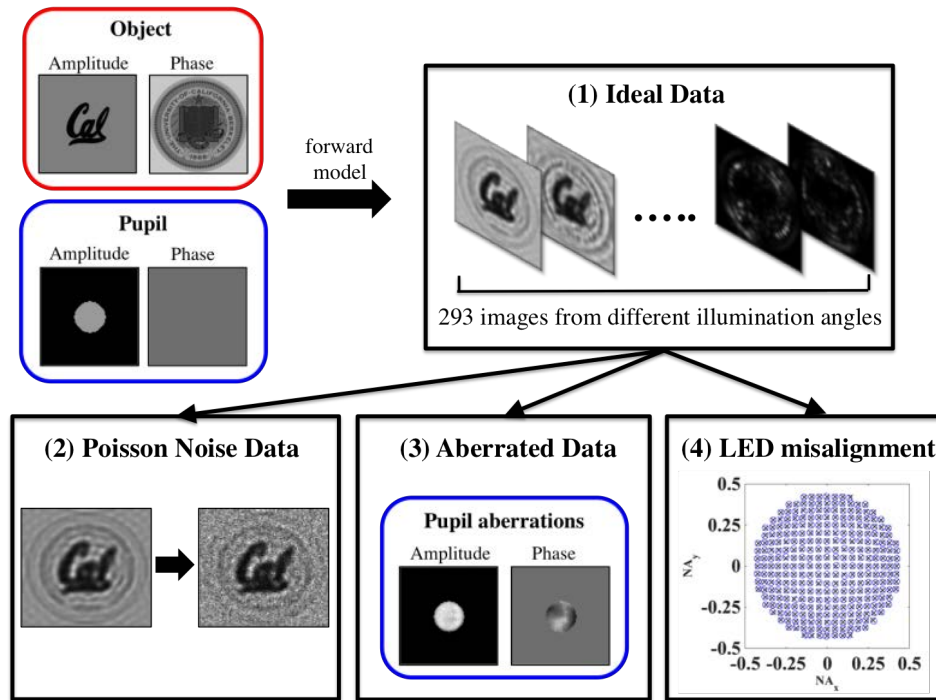


Figure 2.6: To explain the artifacts in our experimental results, as well as evaluate the robustness of various algorithms under common types of errors, we simulate several FPM datasets with different types of known error: (1) Ideal data, (2) Poisson noise data, (3) aberrated data, (4) LED misaligned data (\times : original position, \circ : perturbed position).

4. LED mis-aligned data: the illumination angle of each LED is perturbed slightly (following a normal distribution with standard deviation $\sigma_\theta = 0.2^\circ$). The black \times and blue \circ in Fig. 2.6 show the original and perturbed LED positions, respectively.

2.4.2 Reconstruction results from simulated and experimental dataset

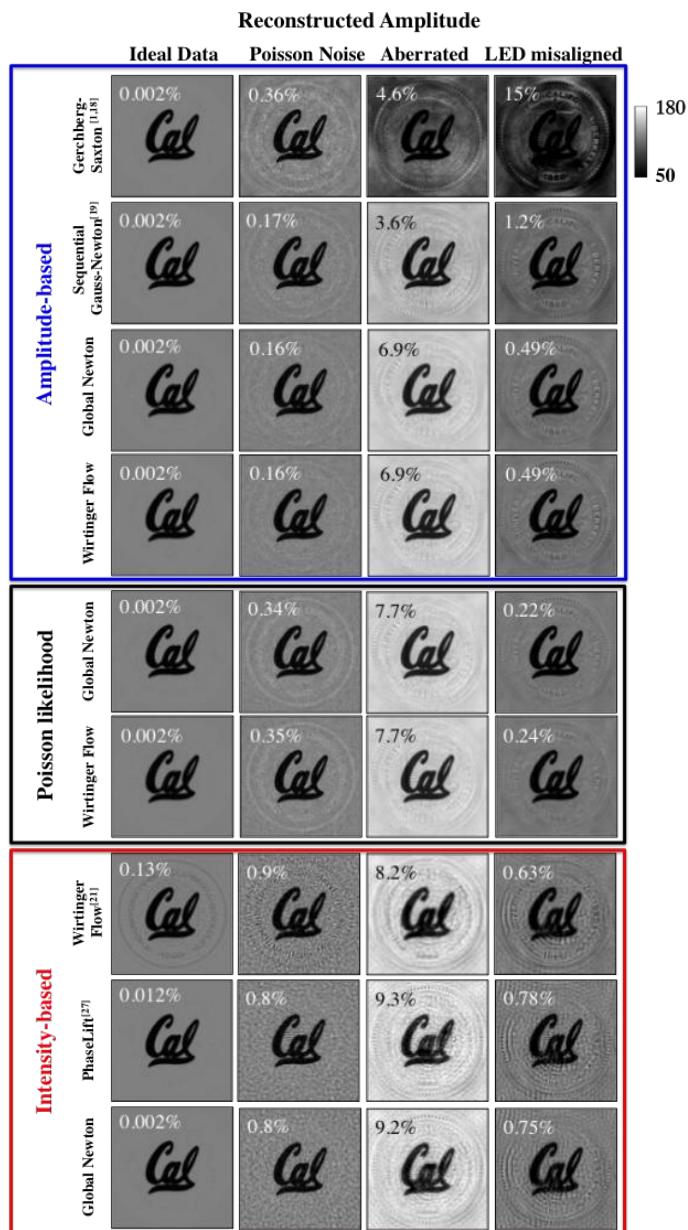


Figure 2.7: Reconstructed amplitude from simulated datasets with three types of errors, using different algorithms. The intensity-based algorithms suffer from high frequency artifacts under both noise and model mis-match errors. The percentage on the top left corner of each image is the relative error of each reconstruction.

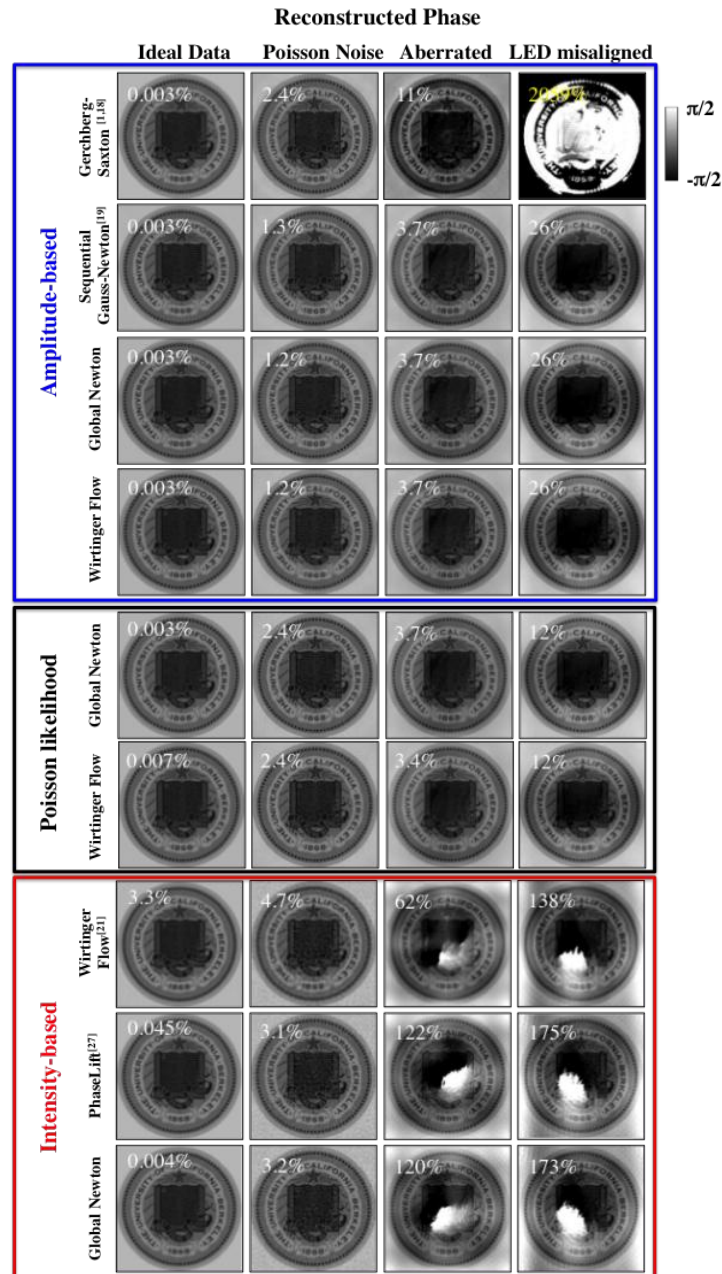


Figure 2.8: Reconstructed phase from simulated datasets with three types of errors, using different algorithms. The intensity-based algorithms suffer from phase wrapping artifacts under both noise and model mis-match errors. The percentage on the top left corner of each image is the relative error of each reconstruction.

Next, we use each of the algorithms described in Section 2.3 to reconstruct amplitude and phase from the datasets simulated in Section 2.4.1, in order to quantify performance under

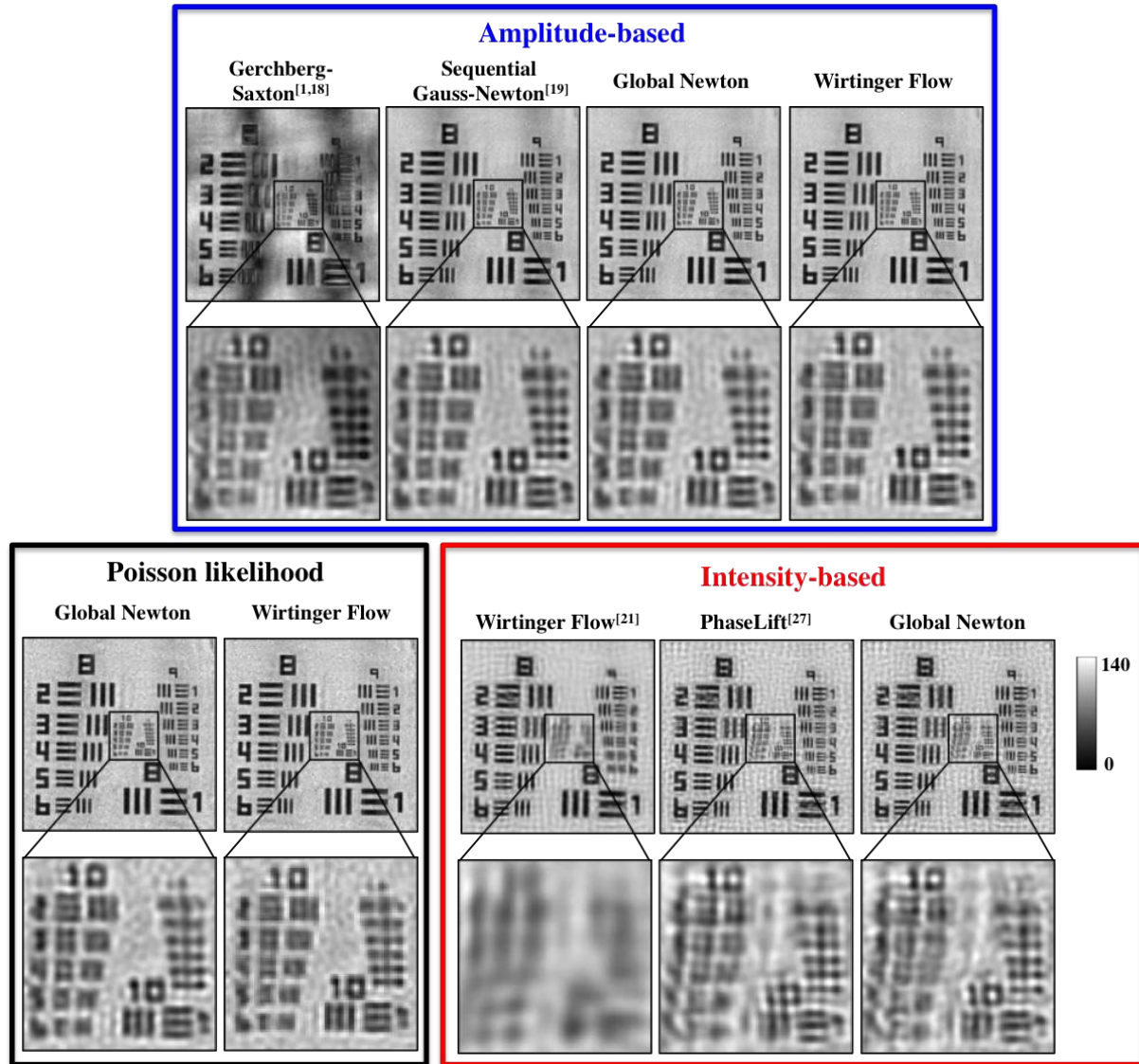


Figure 2.9: Fourier ptychographic reconstruction (amplitude only) of a test object with the algorithms discussed here, all using the same experimental dataset. Algorithms derived from the same cost function (amplitude-based, intensity-based, and Poisson-likelihood) give similar performance, and first-order methods (Gerchberg-Saxton) suffer artifacts.

various experimental error types by comparing against the ground truth input. Figures 2.7 and 2.8 show the reconstructed amplitude and phase, respectively. On the top left corner of each image we give the relative error of the reconstruction, defined as

$$\text{Error} = \frac{\|\mathbf{O}_{\text{recover}} - \mathbf{O}_{\text{true}}\|_2^2}{\|\mathbf{O}_{\text{true}}\|_2^2}, \quad (2.44)$$

where $\mathbf{O}_{\text{recover}}$ and \mathbf{O}_{true} are the reconstructed and true images, respectively, in vector form. In order to ensure that all algorithms converge to their stable solutions, we use 200 iterations for each algorithm, except for Wirtinger flow, which requires 500 iterations. The tuning parameters for each algorithm are summarized in Table 2.1. We have attempted to optimize each parameter as fairly as possible; for example, we use a large σ in the PhaseLift algorithm to achieve a better reconstruction. Small σ trades resolution for flatter background artifacts.

Table 2.1: Tuning Parameters

Gerchberg Saxton	Sequential Gauss-Newton	Amplitude Newton	Amplitude Wirtinger	Poisson Newton	Poisson Wirtinger
N/A	$\delta = 5$	N/A	$i_0 = 10$ $\theta_{\max} = 0.05$	N/A	$i_0 = 10$ $\theta_{\max} = 0.05$
Intensity Wirtinger	PhaseLift	Intensity Newton			
$i_0 = 10$ $\theta_{\max} = 1$	$\sigma = 10^{10}$	N/A			

In analyzing results from the simulated datasets, we find that algorithms with the same cost function give similar reconstruction artifacts. For example, the *intensity-based* algorithms suffer from high-frequency artifacts and phase wrapping when the data is not perfect. Almost all algorithms give a satisfactory reconstruction when using the error-free ideal dataset, except for *intensity-based* Wirtinger flow, which suffers some phase-amplitude leakage and phase blurring (see Figs. 2.7-2.8). When the dataset contains noise or model mis-match, we observe a distinct trend that *amplitude-based* and *Poisson-likelihood-based* algorithms give a better result, compared with *intensity-based* algorithms. The exception to this trend is the Gerchberg-Saxton algorithm, which is somewhat unstable and gets stuck in local minima, so is not robust to any type of error.

The goal of our simulations was to determine the main error sources that cause artifacts in the experimental reconstructions of Fig. 2.9. Since the experiments contain combined errors from multiple sources, it is difficult to attribute artifacts to any particular type of error. We find, however, that all three of our main error sources cause similar artifacts, hence our experimental results may be corrupted by any of Poisson noise, aberration, or LED misalignment. For example, notice that our simulated error-corrupted data all results in high-frequency artifacts when using *intensity-based* algorithms, similar to the experimental results. The Gerchberg-Saxton result also displays low-frequency errors in simulation, as in experiment. The fact that both noise and model mis-match create similar artifacts is unexpected, since they are very different error mechanisms. We explain below why all three are intensity-dependent errors, which is the reason why the cost function choice is so important for robustness. The consequence is that algorithms which use a more accurate noise model (amplitude and Poisson likelihood-based) will not only be more robust to noise, but also to model mis-match errors.

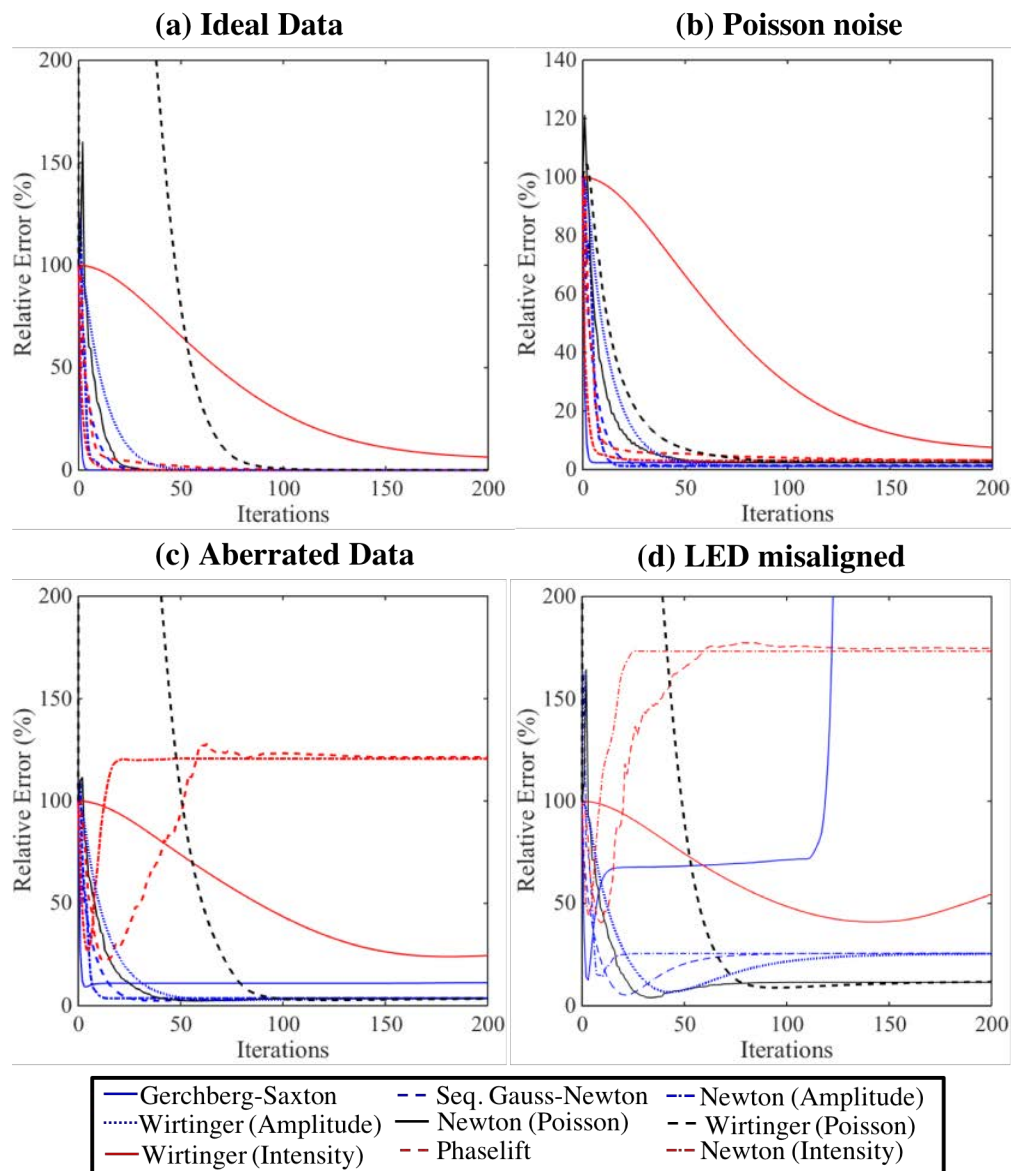


Figure 2.10: Phase relative error as a function of iteration number for different algorithms with the (a) ideal data, (b) Poisson noise data, (c) aberrated data and (d) LED misaligned data. When the data is not perfect, some of the algorithms may not converge to a correct solution.

To examine the convergence of each algorithm, Figure 2.10 plots the error for each iteration when using the aberrated dataset and LED misaligned dataset with different algorithms. The *intensity-based* algorithms (red curves) clearly do not converge to the correct solution and can incur large errors when the data is not perfect. Compared to Phaselift and the intensity-based Newton’s method, the Wirtinger-flow algorithm seems to have lower error;

however, this is only due to its slow convergence. If run it for many iterations, it will eventually settle on a similarly error-corrupted result as the other two *intensity-based* algorithms (not shown). We also observe that *amplitude-based* (blue curves) and *Poisson-likelihood-based* (black curve) algorithms converge to points with lower errors in a similar fashion. This behavior is well explained by the similarity of the algorithms in their use of gradients and Hessians (as shown in the Appendix). Again, the exception to the trend is the first-order Gerchberg-Saxton algorithm, which recovers the object fairly well with aberrated data, but goes unstable in the case of LED misalignment. Note that, when there is no pupil estimation step, the only difference between the Gerchberg-Saxton and the sequential Gauss-Newton algorithm is the step size. Since the latter algorithm gives a good reconstruction, while the former diverges, we conclude that the Gerchberg-Saxton step size is too large for a stable update in this particular case.

Table 2.2: Convergence Speed

	Ideal data		Misaligned data	
	Iteration number	Runtime (s)	Iteration number	Runtime (s)
Gerchberg Saxton	4	2.22	diverges	diverges
Sequential Gauss-Newton	23	12.97	83	46.8
Amplitude Newton	13	100.49	20	154.6
Amplitude Wirtinger	46	26.28	158	89.52
Poisson Newton	28	211.68	77	582.1
Poisson Wirtinger	96	54.46	153	87.36
Intensity Wirtinger	1481	651.64	diverges	diverges
PhaseLift	67	386.28	diverges	diverges
Intensity Newton	12	74.44	diverges	diverges

The convergence speed of each algorithm can be determined from Figure 2.10 using two metrics: *number of iterations* required and total *runtime*. We choose the convergence curves from the cases of ideal data and LED misaligned data and compare their iteration numbers and runtimes in Table 2.2. All the algorithms were implemented in MATLAB on an Intel i7 2.8 GHz CPU computer with 16G DDR3 RAM under OS X operating system. We define convergence as the point when the relative phase error reaches its stable point. The comparison does not consider the divergent cases. In the ideal data case, we can see

that the sequential methods outperform all the other algorithms in terms of runtime. The Gerchberg-Saxton algorithm is the fastest in terms of both iteration number and runtime for this perfect dataset. The global Newton's method using *intensity-based* and *amplitude-based* cost functions also converge very fast in terms of iteration number. The Wirtinger flow algorithm takes much longer to reach convergence both in runtime and iteration number. For the case of the LED misaligned data, only five algorithms converge. In terms of iteration number, the *amplitude-based* Newton's method converges much faster than the other four, as expected. However, the sequential Gauss-Newton algorithm converges much faster in terms of the runtime. Though the global Newton's method is theoretically better than the others, it takes significant time to calculate the full Hessian matrix. Thus, the sequential Gauss-Newton method is our preferred algorithm in practice, because it provides excellent robustness while also enabling fast runtimes and reasonable computational complexity.

The main conclusions to be drawn from this section are that the FPM optimization algorithms which are formulated from *amplitude-based* and *Poisson-likelihood-based* cost functions are more tolerant to imperfect datasets with both Poisson noise and physical deviations like model mis-match, which were represented by aberrations and LED misalignment here. In the next section, we will explain more about the causes for this trend.

2.4.3 Noise model analysis

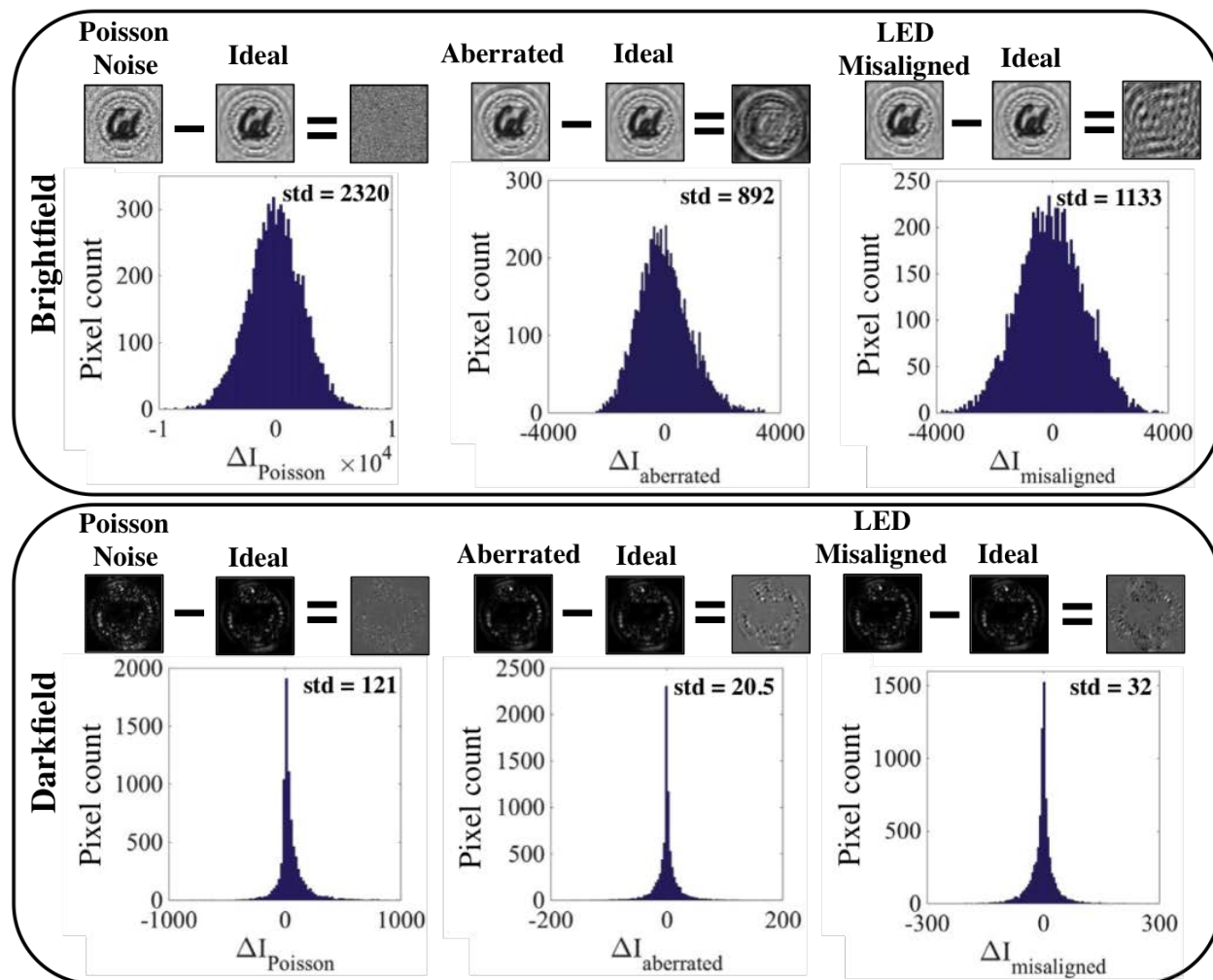


Figure 2.11: Both Poisson noise and model mis-match (aberrations, LED misalignment) cause errors that scale with mean intensity. Here, histograms show the intensity deviations under Poisson noise, aberration, and misalignment for a brightfield and darkfield image.

The reason why *amplitude-based* and *Poisson-likelihood-based* algorithms have superior tolerance to experimental errors is due to their Poisson noise model. Each of these algorithms makes an implicit or explicit assumption that the magnitude of the errors in the data scale with the measured intensity. This is obviously a good model for Poisson noise errors, which are defined as noise which scales with intensity. It is not as obvious that the model mis-match errors (aberrations and LED misalignment) scale with intensity as well. To demonstrate this, Fig. 2.11 shows the histogram of the difference between the deviated dataset and the ideal dataset, for the cases of both brightfield and darkfield images. The histograms show a similar

trend - all of the brightfield errors are much larger than the darkfield errors, with a similar statistical variation. Thus, the errors from Poisson noise, aberrations *and* LED misalignment all scale with the measured intensity. In our experimental data, there are always aberrations in the objective lens, LED misalignment, and Poisson shot noise. Since the noise model for the *amplitude-based* and *Poisson-likelihood-based* algorithms match the actual noise properties, these algorithms perform better than the *intensity-based* algorithms. And since the images captured by FPM have drastically different intensity values, this effect dominates the reconstruction artifacts. Note that these large variations in intensity values are specific to FPM and likely do not play a major role in other phase imaging schemes (e.g. phase from defocus or traditional ptychography), where images do not have such a wide range of intensity values. In our experiments, the Poisson noise is fairly low (due to use of a high-performance sCMOS sensor), but the model mismatch in the experimental data can cause effects similar to strong Poisson noise.

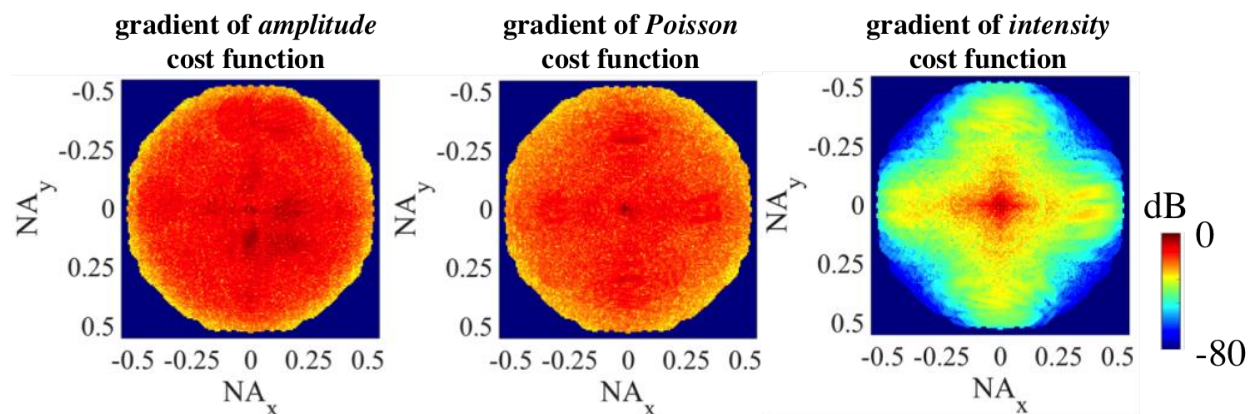


Figure 2.12: The *intensity-based* cost function gives higher weighting to images in the low spatial frequency region of the Fourier domain, resulting in high-frequency artifacts. Here, we show the gradient of the *amplitude-based*, *Poisson-likelihood-based* and *intensity-based* cost functions at the tenth iteration, using experimental data.

For further understanding, we look closer at the relationship between the noise model and the cost function. Our optimization algorithms are derived from three cost functions. Each of the cost functions makes a noise model assumption. The *intensity-based* cost function assumes that noise in the data follows a white Gaussian noise model, which means that the standard deviation of the noise is assumed to be the same across the brightfield and darkfield images. Recall that the standard deviation of a Gaussian noise probability model is related to the weight in the cost function for each pixel, as shown in Eq. 2.4. The larger the standard deviation (amount of noise) at any pixel in Fourier space, the smaller the weighting, since noisy pixels should be trusted less. In the Gaussian noise model, the weights in the cost function for large-valued pixels and small-value pixels are the same. However, the deviation for brightfield images is much larger than that for darkfield images, as shown in Fig. 2.11.

Hence, the brightfield images will contribute more to the total cost function value if the weights are all the same, due to their high intensity. The result is that the *intensity-based* (Gaussian noise model) algorithms focus mostly on the brightfield images, which correspond to low spatial frequency information, and the darkfield images do not contribute much. The result is a failure in the high-frequency reconstruction, as we saw in Figs. 2.9, 2.7, 2.8, and loss of effective resolution since the darkfield images contain all the sub-diffraction-limit information. To illustrate the dramatic difference in weights, Fig. 2.12 shows the gradient of the different cost functions. Obviously, the intensity cost function gives much higher weighting to low spatial frequencies, which causes the high-frequency artifacts.

Since the amplitude-based cost function shares a similar gradient and Hessian with the Poisson likelihood function, as shown in the Appendix and Fig. 2.12, it is not surprising that they both produce a similar quality reconstruction. Both of these cost functions assume the noise in the data follow a Poisson distribution, with the standard deviation scaling with the measured intensity. This assumption matches the actual error better than the white Gaussian assumption. The actual noise or deviations in the experiments for brightfield images have larger standard deviation, while that for darkfield images have smaller standard deviation. Under the Poisson noise model, the weight in the cost function is smaller for the noisy brightfield images and larger for the darkfield images. At the end, algorithms based on the Poisson noise model put more emphasis on the darkfield images and thus get a better reconstruction compared to the *intensity-based* algorithms. Figure 2.12 shows that the gradients for the *amplitude-based* and *Poisson-likelihood-based* cost function are similar and are more uniform throughout the whole Fourier space.

2.5 Joint estimation of pupil function and LED positions

In the previous section, we have shown that model mis-matches such as aberrations and LED mis-alignment greatly affect the final reconstruction no matter what kinds of algorithms are used. Thus, it is important to develop a sub-routine in the main algorithm that can correct for these mis-matches. In this section, we will describe how we correct for these mis-matches by including sub-optimization steps within the main algorithms. The flow to correct for these mis-matches are summarized in Fig. 2.13.

2.5.1 Pupil recovery

There are already more sophisticated FPM extensions to correct for some model mis-match errors [21, 93], similar to the probe correction algorithms in traditional ptychography [91]. Both of the methods previously developed for Fourier ptychography are derived from the amplitude-based formulation. By taking the derivative of the cost function with respect to

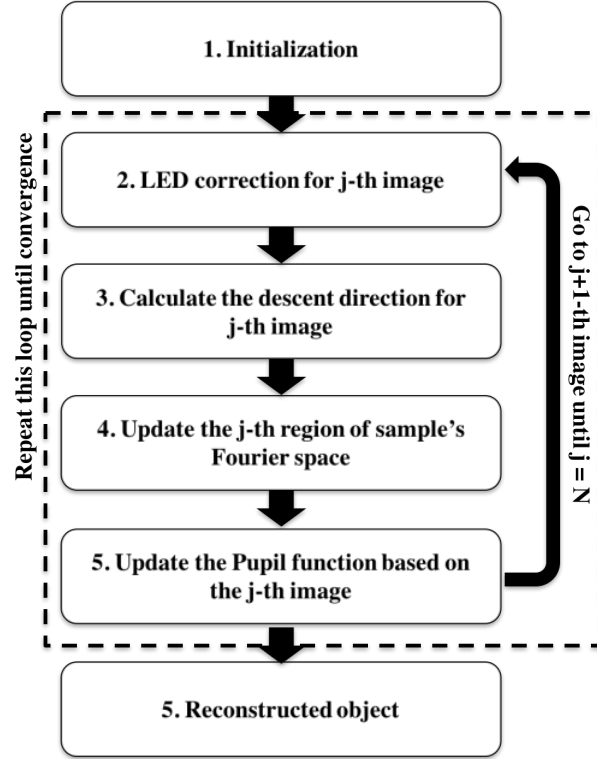


Figure 2.13: The flow of model mis-match correction algorithms for Fourier ptychographic microscopy. Redundancy and diversity in the dataset enables algorithmic self-calibration of aberrations (pupil function) and LED position errors.

\mathbf{P} , the decent direction to estimate the pupil function can be calculated as

$$\nabla_{\mathbf{P}} f_{A,\ell}(\mathbf{O}, \mathbf{P}) = -\text{diag}(\bar{\mathbf{Q}}_{\ell} \bar{\mathbf{O}}) \left[\mathbf{F} \text{diag} \left(\frac{\sqrt{\mathbf{I}_{\ell}}}{|\mathbf{g}_{\ell}|} \right) \mathbf{g}_{\ell} - \text{diag}(\mathbf{P}) \mathbf{Q}_{\ell} \mathbf{O} \right]. \quad (2.45)$$

By applying the pupil estimation step after each object estimation using this gradient or approximated Hessian, the sequential gradient descent [93] and the sequential Gauss-Newton method [21] including pupil estimation can be derived. Here we only consider the amplitude-based cost function, for simplicity.

We wish to investigate the improvements obtained by adding a pupil estimation step to both first and second-order optimization algorithms. Figure 2.14 shows the reconstruction result from the sequential gradient descent (first-order) and sequential Gauss-Newton (second-order) algorithms, using the aberrated dataset from the previous simulations. The numbers at the top left corner are the relative error compared to the ground truth simulated image. As can be seen, adding the pupil estimation step gives a better complex-field reconstruction, and the second-order (Gauss-Newton) method with pupil estimation provides the best result.

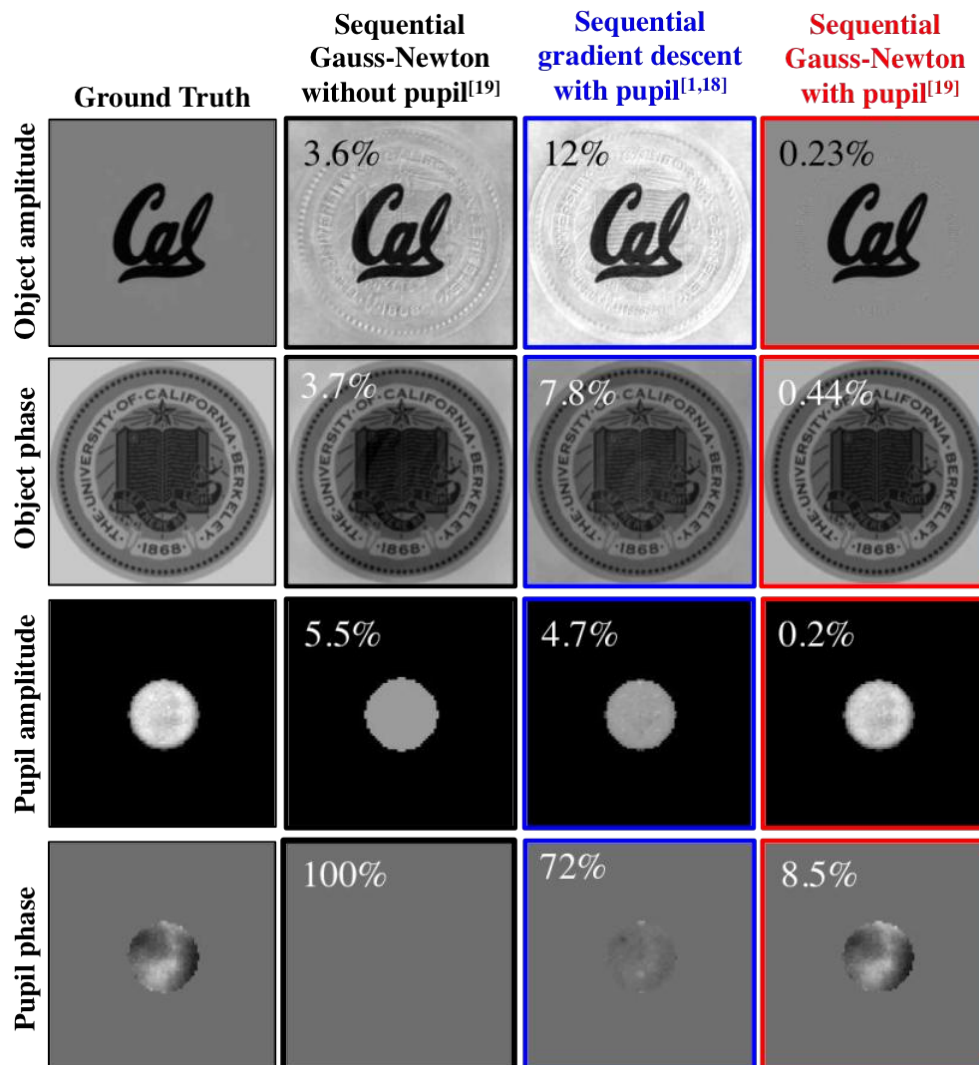


Figure 2.14: Object and pupil reconstruction results using different algorithms, with and without pupil estimation. The second-order method (sequential Gauss-Newton) with pupil estimation gives the best result, as expected. In this case, we find that the second-order method *without* pupil estimation is already better than first-order method (sequential gradient descent) *with* pupil estimation.

Surprisingly, however, the second-order reconstruction without pupil estimation is better than the first-order reconstruction with pupil estimation, for this case. This highlights the robustness to aberrations that a second-order optimization scheme enables. The second-order nature of the algorithm makes it faster in convergence, and also more stable. In terms of runtime, the pupil estimation step takes about the same time as the object reconstruction part, so the algorithm is two times slower when the pupil function step is incorporated.

2.5.2 LED position recovery

Another possible correction scheme for model mis-match is that for LED misalignment. Since each LED position corresponds to a certain shift of the pupil function in the Fourier domain, this is similar to the shift of the probe function in traditional ptychography. There, iterative algorithms have been proposed to correct for the positioning error of the probe function [89, 105–107]. In [89, 107], a gradient of the cost function with respect to the shift of the probe function has been calculated and the conjugate gradient method has been applied to correct for the positioning error. In [105], a simulated annealing method is adopted to estimate the shift of the probe function. The simulated annealing method is also adopted to correct for the misalignment of the spatial light modulator in a overlapped Fourier coding system [108], analogous to FPM. In our experiments, we observe that the simulated annealing method can locate the LED positions more accurately than other methods. Thus, we only compare with the simulated annealing method.

2.5.2.1 Simulated annealing position estimation

Simulated annealing is a method of searching unknown variables over a finite space to minimize or maximize the function of merit - the cost function in our case. Instead of exhaustively testing all the possible states, simulated annealing iteratively approaches the optimal state. At the first iteration, the algorithm randomly searches several states in the space and selects the one with the smallest cost function value. The algorithm then starts at this state for the next iteration, slowly reducing the search range in the following iterations until convergence.

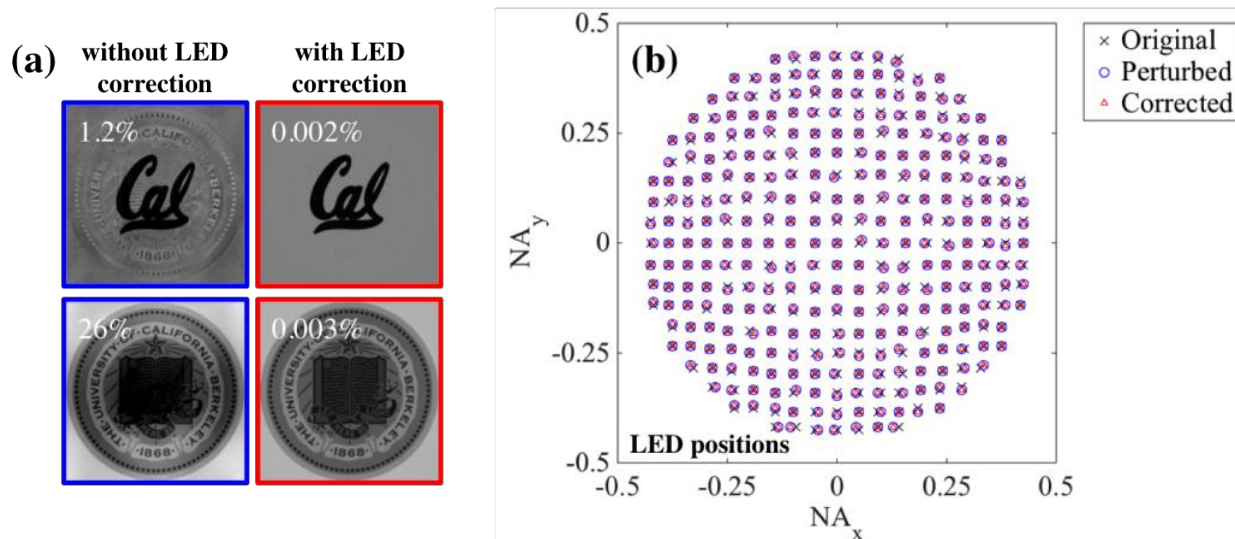


Figure 2.15: (a) Adding LED misalignment correction improves the reconstruction results (sequential Gauss-Newton method). (b) The original, perturbed, and corrected LED positions in angular coordinates. LED correction accurately retrieves the actual LED positions.

In our sequential algorithm, the whole optimization problem is divided into many sub-optimization problems for different collected images. At each sub-optimization problem, a gradient descent or Gauss-Newton method is applied to update that corresponding region in Fourier domain. To add a LED mis-alignment correction step, the simulated annealing algorithm can be incorporated into each sub-iteration to find an optimal shift of the pupil function. In each sub-iteration, the down-sampling matrix, \mathbf{Q}_ℓ , which contains the information of the pupil shift, is tested according to the annealing process for several possible states corresponding to different shifts of the pupil. The state with the smallest cost function value is selected to update the old down-sampling matrix. Then, the new down-sampling matrix is used to update the corresponding region in the Fourier domain.

The simulated annealing method estimates the LED positions with good accuracy. Figure 2.15 shows the reconstruction result from the simulated LED misaligned dataset, both with and without the LED correction step. The result using the LED correction clearly shows better quality and smaller error, as seen in Fig. 2.15(a). Since the LED correction scheme also estimates the actual LED positions, which we intentionally perturbed in order to impose a known error, we can also compare the actual and recovered LED positions, shown in Fig. 2.15(b).

To complete the picture, we now show experimental reconstructions with and without the two correction schemes: pupil correction and LED mis-alignment corrections (see Fig. 2.16). Since we do not know ground truth for our experiments, we can only make qualitative observations. An incremental improvement is observed when adding the pupil estimation and then the LED correction steps - the background variation becomes flatter. Figure 2.16(b) shows the corrected LED positions compared to the original ones, in angular coordinates. Corrected positions of LEDs in different regions share similar offset because the fabrication process of the LED array can cause unexpected position misalignment for each LED. Notice that the LEDs at the edges (corresponding to higher angles of illumination) incur more variation, since these are more sensitive to calibration. Also, many of the large deviations occur at the edges that are not along the horizontal and vertical axes. In these areas, the LED position recovery is poor because the object has very little information there (the resolution test target contains only square features) and so the data contains little information about these areas. However, any errors in LED positions in this area will also not significantly affect the reconstruction if they do not contribute much energy to the object spectrum. If the goal was not to correct the image results, but rather to find the LED positions accurately, then one should choose an object that contains uniformly distributed spatial frequencies (e.g. a random diffuser or speckle field). Although the simulated annealing further improves our reconstruction, we note that it is more than ten times slower to process the data because of the local search performed at each sub-iteration.

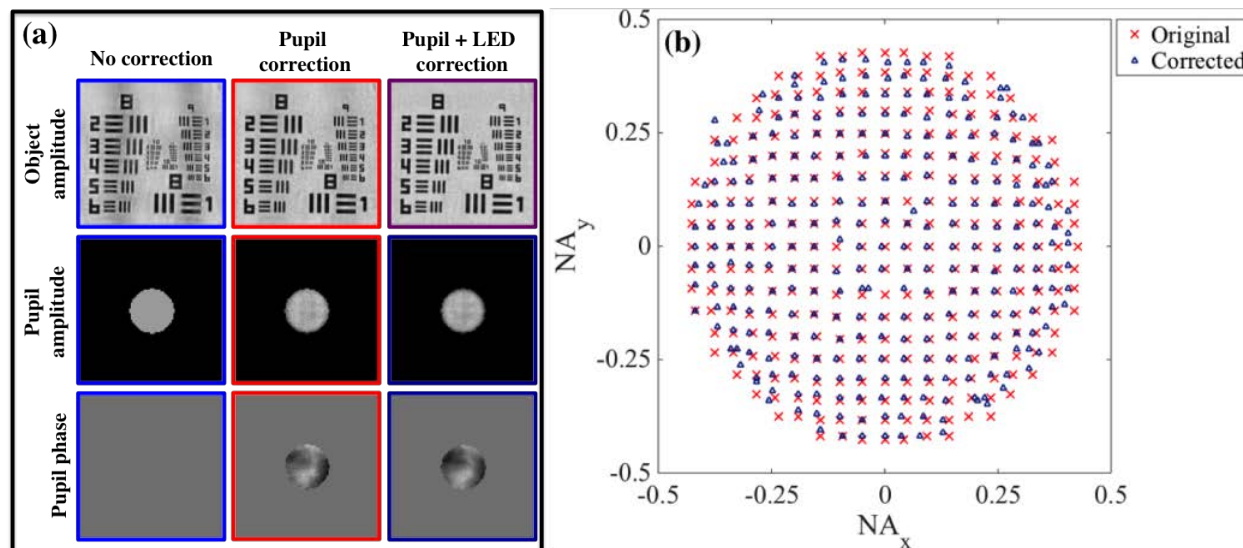


Figure 2.16: Experimental reconstructions with and without LED misalignment correction (sequential Gauss-Newton method). (a) The reconstructed object and pupil. (b) The original and corrected LED positions, in angular coordinates.

2.6 Conclusion

In this chapter, we formulated the Fourier ptychographic phase retrieval problem using maximum likelihood optimization theory. Under this framework, we reviewed the existing FPM algorithms and classified them based on their cost functions as: *amplitude-based* algorithms (akin to a Poisson noise model) and *intensity-based* algorithms (akin to a white Gaussian noise model). We derived a new algorithm based on the Poisson likelihood function, which is more robust to measurement imperfections. We compared the tolerance of these algorithms under errors due to experimental noise and model mis-match (aberrations and LED mis-alignment) using both simulated data and experimental data. Because the noise and model mis-match error for brightfield and darkfield images depend on the measured intensity, the *amplitude-based* and *Poisson-likelihood-based* algorithms that use a the Poisson noise model are more robust than the *intensity-based* algorithms. This can be explained by the standard deviation of the noise model determining the weight of each image in the optimization. Hence, *intensity-based* algorithms over-weight the brightfield images, resulting in poor high-frequency reconstruction, which is where the high-resolution details reside.

Next, we used existing pupil estimation algorithms and proposed a simulated-annealing-based LED correction algorithm for algorithmic self-calibration of model mis-match. We compared the performance of the pupil estimation algorithms and found that second-order methods give the best results. We also showed the capability of the simulated annealing method to correct for misaligned LEDs and find their actual positions.

Based on our studies, we conclude that the global Newton's method gives the best recon-

struction, but with high computational cost. Considering both robustness and computational efficiency, we find that the sequential Gauss-Newton method provides the best trade-offs for large-scale applications. Its experimental robustness is verified in our recent time-series *in vitro* experiments [22], where we demonstrate sub-second acquisition times for FPM.

Chapter 3

Computational random structured illumination fluorescence microscopy

In Chapter 2, we have demonstrated the effectiveness of computational imaging approach to achieve super-resolution coherent imaging. By adopting the optimization formulation, the information retrieval process is no longer restricted to analytic inversion, which makes the optical design more flexible (non-interferometric system for phase imaging). In addition, we also demonstrate the effectiveness of self-calibration algorithm to correct for experimental errors, such as system aberration and illumination mis-alignment. Better image quality could be achieved with the power of computational imaging.

Carrying this idea in mind, in this chapter, we investigate if computational imaging approach could also be a benefit to super-resolution fluorescent imaging system. Structured illumination microscopy (SIM) is also an illumination based super-resolution imaging technique similar to FPM, but for fluorescent microscopy. It improves resolution by down-modulating high-frequency information of an object to fit within the passband of the imaging system as we explained in Chapter 1. Generally, the reconstruction process requires prior knowledge of the illumination patterns, which implies a well-calibrated and aberration-free system. With the help of optimization formulation, we propose a new *algorithmic self-calibration* strategy for SIM that does not need to know the exact patterns *a priori*, but only their covariance. The algorithm, termed PE-SIMS, includes a Pattern-Estimation (PE) step requiring the uniformity of the sum of the illumination patterns and a SIM reconstruction procedure using a Statistical prior (SIMS). Additionally, we perform a pixel reassignment process (SIMS-PR) to enhance the reconstruction quality. We achieve $2\times$ better resolution than a conventional widefield microscope, while remaining insensitive to aberration-induced pattern distortion and robust against parameter tuning. In the end, we also propose a generalization of this method to 3D super-resolution fluorescent imaging.

3.1 Introduction

3.1.1 Short review of structured illumination microscopy

SIM [2, 4–6], as explained in Chapter 1, uses illumination of multiple structured patterns to down-modulate high spatial frequency information of the object into the low-frequency region, which can then pass through the bandwidth of the microscope’s optical transfer function (OTF) and be captured by the sensor. The reconstruction algorithm for SIM combines demodulation process which brings the high spatial frequency information back to its original position and synthetic aperture that extends the support of the effective OTF. Various types of structured patterns have been used to realize SIM: periodic gratings [2, 4–6], a single focal spot (confocal microscope) [43, 44], multifocal spots [45–48] and random speckles [48–57]. When the illumination patterns themselves are diffraction-limited, linear SIM is restricted to $2\times$ the bandwidth of a widefield microscope [6], allowing up to $\sim 2.4\times$ resolution enhancement (metrics explained in Sec. 3.3).

3.1.2 Blind structured illumination microscopy

In practice, structured illumination systems are sensitive to aberrations and experimental errors. To avoid reconstruction artifacts that degrade resolution, the patterns that are projected onto the sample must be known accurately. Periodic grating patterns can be parameterized by their contrast, period and phase angle, which may be estimated in the post-processing [109–112]. For multifocal patterns, the location of each focal spot is required [45]. For random speckle patterns, the relative shifts of the patterns are needed [53, 54]. Even with careful calibration and high-quality optics, distortions caused by the sample may degrade the result.

To alleviate some of the experimental challenges, blind SIM was proposed, enabling SIM reconstruction without many priors [51, 52, 56, 57, 113, 114]. The only assumption is that the sum of all illumination patterns is uniform. Optimization-based algorithms have been adopted, including iterative least squares with positivity and equality constraints [51, 56, 113], joint support recovery [52] and ℓ_1 sparsity constraints [57]. However, these algorithms are sensitive to parameter tuning and may show low contrast in reconstructing high spatial frequencies [51]. Another algorithm, speckle super-resolution optical fluctuation imaging (S-SOFI) realizes SOFI [115] by first projecting random speckle patterns onto the object, and then using the statistical properties of the speckle patterns as a prior to reconstruct a high-resolution image [55]. S-SOFI is experimentally simple and robust; however it only achieves a $1.6\times$ resolution enhancement instead of $2.4\times$ for conventional SIM techniques (as compared to a widefield microscope).

3.1.3 Contribution of this chapter

In this chapter, we propose a new reconstruction algorithm for SIM that is applicable to any illumination patterns. Our method, termed pattern estimation structured illumination microscopy with a statistical prior (PE-SIMS), is as robust and insensitive to parameter tuning as S-SOFI, and achieves better resolution enhancement (up to $2\times$). Like blind SIM, the patterns need not be known (except for a requirement on the covariance of the patterns). We demonstrate our method using simulated and experimental results with both speckle and multifocal patterns. We discuss pattern design strategies to reduce the amount of data required and demonstrate an extension that uses pixel reassignment [116–120] to improve the reconstruction quality. In the end, we generalize this idea to 3D imaging.

3.2 Theory and method

Our algorithm takes in a SIM dataset consisting of multiple images captured under different structured illumination patterns (e.g. random speckles, multifocal spots). We reconstruct the super-resolved image in two parts. The first part is an iterative optimization procedure for estimating each illumination pattern based on an approximated object. The second part reconstructs the high-resolution image using the estimated patterns and the measured images, along with a statistical prior. Before introducing these two parts, we start by defining the SIM forward model.

3.2.1 Forward model of structured illumination microscopy

A representative experimental setup is shown in Fig. 3.1. A DMD spatial light modulator (SLM) is used to project patterns onto the object through an objective lens. The measured intensity for the ℓ -th captured image is the product of the object’s fluorescence distribution $o(\mathbf{r})$ with the illumination pattern $p_\ell(\mathbf{r})$, where $\mathbf{r} = (x, y)$ denotes the lateral position coordinates. This product is then convolved with the system’s incoherent point spread function (PSF) on the detection side, $h_{\text{det}}(\mathbf{r})$:

$$I_\ell(\mathbf{r}) = [o(\mathbf{r}) \cdot p_\ell(\mathbf{r})] \otimes h_{\text{det}}(\mathbf{r}) = \iint o(\mathbf{r}') p_\ell(\mathbf{r}') h_{\text{det}}(\mathbf{r} - \mathbf{r}') d^2\mathbf{r}'. \quad (3.1)$$

3.2.2 Part 1: Pattern estimation

The first part of our inverse algorithm is to estimate the illumination patterns. To do so, we start with a low-resolution approximation of the object. Then, we use this object and our measured images to iteratively estimate the patterns (see Fig. 3.2).

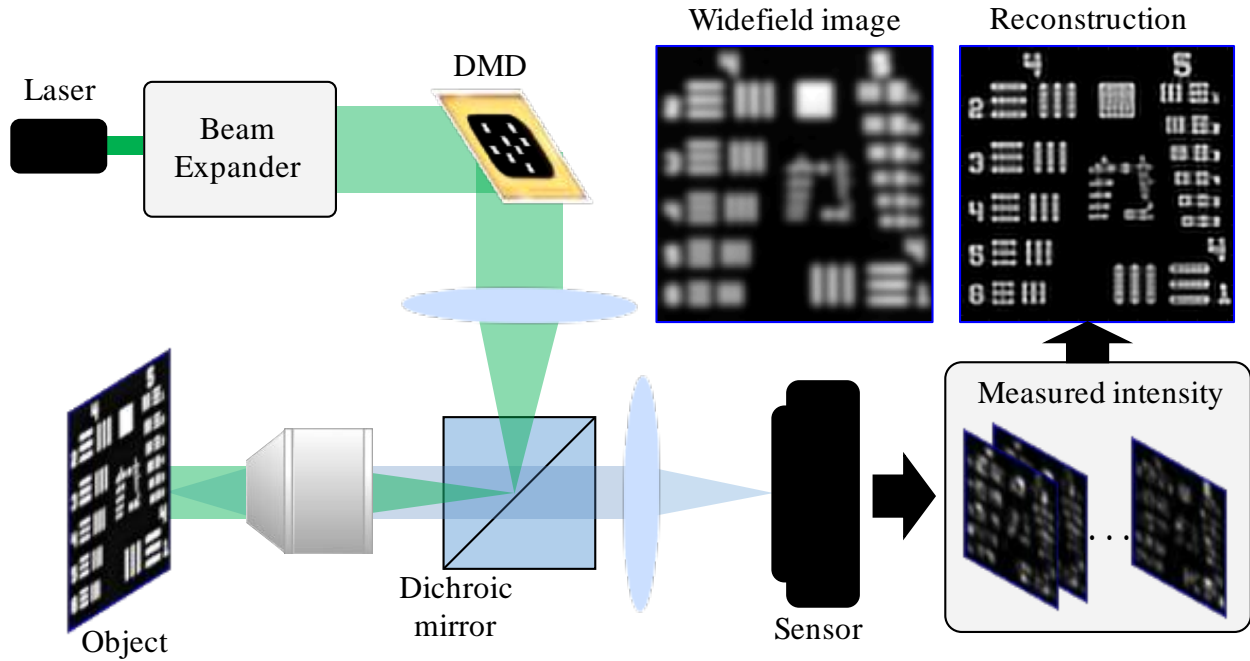


Figure 3.1: Example experimental setup for structured illumination microscopy (SIM) using a deformable mirror device (DMD) to capture low-resolution images of the object modulated by different illumination patterns. Our IPE-SIMS algorithm reconstructs both the super-resolved image and the unknown arbitrary illumination patterns.

Part 1a: Approximate widefield image

If we already knew the object $o(\mathbf{r})$, it would be straightforward to estimate the pattern for each measured image by dividing out the object from each of the measurements. However, the object $o(\mathbf{r})$ is unknown. Hence, we start by making a rough estimate of the object. We first take the mean of all the measured images:

$$I_{\text{avg}}(\mathbf{r}) = \langle I_\ell(\mathbf{r}) \rangle_\ell = [o(\mathbf{r}) \cdot \langle p_\ell(\mathbf{r}) \rangle_\ell] \otimes h_{\text{det}}(\mathbf{r}) \approx p_0 o(\mathbf{r}) \otimes h_{\text{det}}(\mathbf{r}), \quad (3.2)$$

where $\langle \cdot \rangle_\ell$ is the mean operation with respect to ℓ , and $p_0 = \langle p_\ell(\mathbf{r}) \rangle_\ell$ is approximately a constant over the entire field of view. The resulting image will be equivalent to the low-resolution widefield image if the sum of all illumination patterns is approximately uniform.

Part 1b: Deconvolve widefield image

Since the widefield image represents the convolution of the object with its PSF, we can perform a deconvolution operation to estimate the low-resolution object:

$$o_{\text{est}}(\mathbf{r}) = \mathcal{F}^{-1} \left\{ \frac{\tilde{I}_{\text{avg}}(\mathbf{u}) \cdot \tilde{h}_{\text{det}}(\mathbf{u})}{|\tilde{h}_{\text{det}}(\mathbf{u})|^2 + \beta} \right\}, \quad (3.3)$$

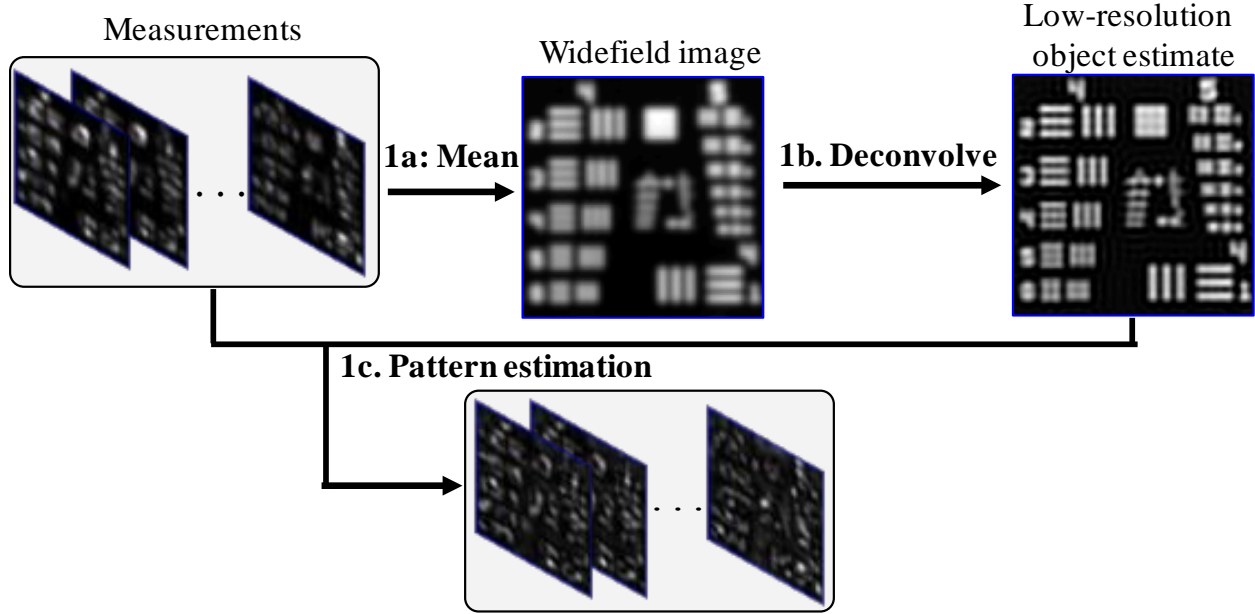


Figure 3.2: The first part of our algorithm, Pattern Estimation (PE), iteratively estimates the illumination patterns from an approximated object given by the deconvolved widefield image.

where \mathcal{F} and \mathcal{F}^{-1} denote the Fourier transform and its inverse, respectively, $\tilde{\cdot}$ denotes the Fourier transform of a certain function, $\mathbf{u} = (u_x, u_y)$ are the lateral spatial frequency coordinates and β is a small Tikhonov regularization constant. Note that this object estimate has diffraction-limited resolution and will be used only for estimating the illumination patterns.

Part 1c: Pattern estimation

We then use the low-resolution object estimate $o_{\text{est}}(\mathbf{r})$ to recover each of the illumination patterns. Since each image is simply the product of the illumination and object, we could divide each image by the estimated object to get the pattern. However, we instead solve the problem as an optimization procedure in order to impose the correct Fourier support constraint and avoid reconstruction artifacts. The ℓ -th pattern estimate is the solution to the following problem

$$\begin{aligned} \underset{p_\ell}{\text{minimize}} \quad & f(p_\ell) = f_{\text{diff}}(p_\ell) + \mathbb{I}_{\mathcal{C}}(p_\ell) = \sum_{\mathbf{r}} |I_\ell(\mathbf{r}) - [o_{\text{est}}(\mathbf{r}) \cdot p_\ell(\mathbf{r})] \otimes h_{\text{det}}(\mathbf{r})|^2 + \mathbb{I}_{\mathcal{C}}(p_\ell), \\ \text{where} \quad & \mathbb{I}_{\mathcal{C}}(p_\ell) = \begin{cases} 0, & p_\ell \in \mathcal{C} \\ +\infty, & p_\ell \notin \mathcal{C} \end{cases}, \quad \mathcal{C} = \left\{ p_\ell(\mathbf{r}) \left| \tilde{p}_\ell(\mathbf{u}) = 0, \forall \mathbf{u} > \frac{2NA}{\lambda_{\text{illu}}} \right. \right\}, \end{aligned} \quad (3.4)$$

where λ_{illu} is the wavelength of the excitation light. The first term of the cost function, $f_{\text{diff}}(p_\ell)$, in Eq. (3.4) is the least square error (residual) between the measured intensity and

the predicted intensity based on our current estimate. The second term enforces a frequency support constraint for the illumination pattern via an indicator function $\mathbb{I}_{\mathcal{C}}$. This is important to reduce artifacts in the pattern estimation because a normal division between the measured image and estimated object will create errors outside of this frequency support. In our epi-illumination geometry, the constraint is that the frequency content of each illumination pattern be confined within the OTF defined by the objective's NA.

We implement a proximal gradient descent algorithm [121], summarized in Algorithm 1. Proximal gradient descent is designed to solve convex optimization problems like ours that have two cost function terms: one being a differentiable cost function term (e.g. the residual) and the other being a constraint or regularization term (usually nondifferentiable). When the constraint is defined by an indicator function, as in Eq. (3.4), the method is also known as a projected gradient method.

To implement, we first compute the gradient of the differentiable cost function term with respect to $p_{\ell}(\mathbf{r})$

$$g_{\ell}^{(k)}(\mathbf{r}) = \frac{\partial f_{\text{diff}}(p_{\ell}^{(k)})}{\partial p_{\ell}} = -o_{\text{est}}(\mathbf{r}) \cdot [h_{\text{det}}(\mathbf{r}) \otimes (I_{\ell}(\mathbf{r}) - [o_{\text{est}}(\mathbf{r}) \cdot p_{\ell}^{(k)}(\mathbf{r})] \otimes h_{\text{det}}(\mathbf{r}))], \quad (3.5)$$

where k denotes evaluation of the gradient using the pattern at the k -th iteration.

We define the projection operation $\Pi_{\mathcal{C}}$ to force the information outside of the OTF to be zero at each iteration. To reduce high-frequency artifacts, the following soft-edge filter is used

$$\Pi_{\mathcal{C}}(y) = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}\{y\} \cdot |\tilde{h}_{\text{illu}}(\mathbf{u})|^2}{|\tilde{h}_{\text{illu}}(\mathbf{u})|^2 + \delta} \right\}, \quad (3.6)$$

where $h_{\text{illu}}(\mathbf{r})$ is the system's PSF on the illumination side, and δ determines the amount of high-frequency information that is suppressed in the pattern estimation step. We repeat this process of updates and projections until convergence (typically ~ 50 iterations to estimate each pattern).

The convergence speed for proximal gradient descent is on the order of $O(1/K)$ [121], indicating that the residual between the current and optimal cost functions is inversely proportional to the number of iterations K . To accelerate convergence, one extra step is conducted in Algorithm 1 to include the information of the previous estimate [122, 123]. The convergence rate for this accelerated proximal gradient method, $O(1/K^2)$ [123], is significantly faster than the normal proximal gradient method.

Algorithm 1 Pattern Estimation

Require: $I_\ell(\mathbf{r})$, $o_{\text{est}}(\mathbf{r})$

- 1: initialize $p_\ell^{(1)}(\mathbf{r})$ with all zero image $t_1 = 1$
- 2: **for** $k = 1 : K$ **do**
- 3: Select step size $\eta^{(k)} > 0$
- 4: $\hat{p}_\ell^{(k+1)}(\mathbf{r}) = \Pi_{\mathcal{C}} \left[p_\ell^{(k)}(\mathbf{r}) - \eta^{(k)} g_\ell^{(k)}(\mathbf{r}) \right]$, where $\Pi_{\mathcal{C}}$ denotes the projection onto \mathcal{C} .
- 5: $t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}$
- 6: $p_\ell^{(k+1)}(\mathbf{r}) = \hat{p}_\ell^{(k)}(\mathbf{r}) + \frac{t_k - 1}{t_{k+1}} \left[\hat{p}_\ell^{(k+1)}(\mathbf{r}) - \hat{p}_\ell^{(k)}(\mathbf{r}) \right]$
- 7: **end for**
- 8: **Output:** $p_\ell(\mathbf{r})$

3.2.3 Part 2: SIM with a statistical prior

Once we have recovered the illumination patterns, the second part of the algorithm is to reconstruct a high-resolution image from the measured dataset $I_\ell(\mathbf{r})$ and the estimated patterns $p_\ell(\mathbf{r})$. We call this part of the algorithm Structured Illumination Microscopy with a Statistical prior (SIMS), summarized in Fig. 3.3. There are four steps, which are explained below. We will also describe how the statistical prior is used and why this procedure gives better resolution.

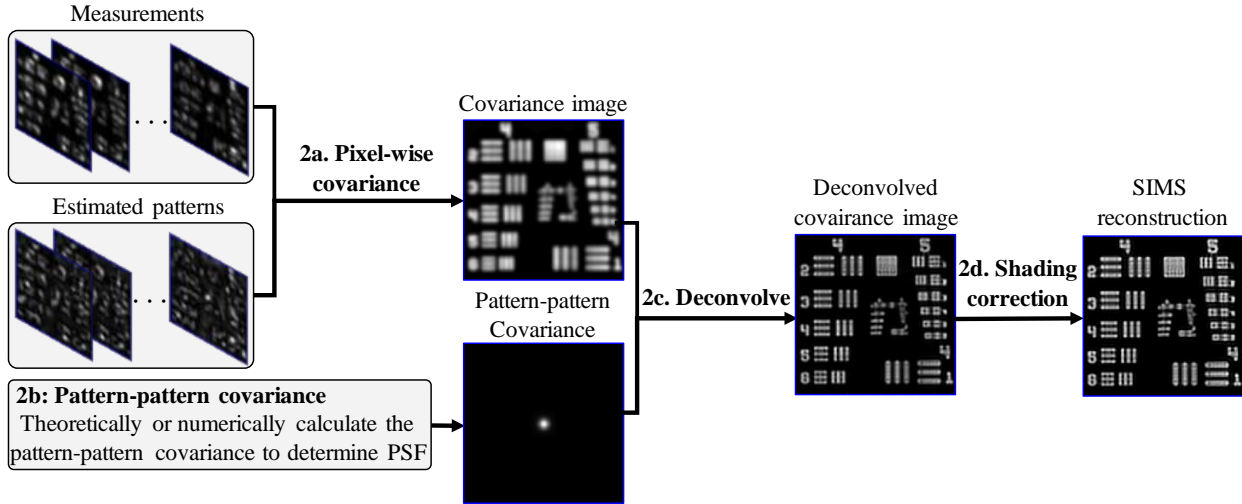


Figure 3.3: The second part of our algorithm, termed structured illumination microscopy with a statistical prior (SIMS), estimates the high-resolution object from the measured images and the estimated illumination patterns obtained in Part 1.

Part 2a: Calculate the pattern-intensity covariance

Consider the case where the pattern $p(\mathbf{r})$ is a random variable at position \mathbf{r} and the measured intensity $I(\mathbf{r})$ is also a random variable at position \mathbf{r} . The ℓ -th image is thus the ℓ -th sample function for these random variables (one event out of the sample space). Covariance is a measure of how much two random variables change together. Since the intensity $I(\mathbf{r})$ is the blurred version of the product between random patterns $p(\mathbf{r})$ and deterministic object $o(\mathbf{r})$ (Eq. (3.1)), the covariance between the pattern and the intensity should give high similarity wherever the object $o(\mathbf{r})$ has signal and thus allow us to find the object underneath the random-pattern illumination [49, 124–127]. We calculate this covariance image $I_{\text{cov}}(\mathbf{r})$ as

$$I_{\text{cov}}(\mathbf{r}) = \langle \Delta p_\ell(\mathbf{r}) \Delta I_\ell(\mathbf{r}) \rangle_\ell = \iint o(\mathbf{r}') \langle \Delta p_\ell(\mathbf{r}) \Delta p_\ell(\mathbf{r}') \rangle_\ell h_{\text{det}}(\mathbf{r} - \mathbf{r}') d^2 \mathbf{r}', \quad (3.7)$$

where $\Delta I_\ell(\mathbf{r}) = I_\ell(\mathbf{r}) - \langle I_\ell(\mathbf{r}) \rangle_\ell$, and $\Delta p_\ell(\mathbf{r}) = p_\ell(\mathbf{r}) - \langle p_\ell(\mathbf{r}) \rangle_\ell$.

Regardless of which illumination pattern is imposed, the covariance image always gives an estimate of the object. However, the resolution of the reconstructed object may be different for different pattern statistics. We can quantify this by taking a closer look at the expression on the right-hand side of Eq. (3.7). The covariance between $p(\mathbf{r})$ and $p(\mathbf{r}')$ acts as the PSF of the covariance image, which thus determines resolution. If the patterns are perfectly spatially correlated, the pattern-pattern covariance is a constant, and the pattern-intensity covariance image is a normal widefield image with PSF of $h(\mathbf{r})$. If the patterns are perfectly spatially uncorrelated, the pattern-pattern covariance is $\langle |\Delta p_\ell(\mathbf{r})|^2 \rangle_\ell \delta(\mathbf{r} - \mathbf{r}')$, which, for a constant variance, results in the PSF being a delta function and the object being reconstructed with perfect resolution. In practice, this is not achievable, since the illumination is bandlimited and thus cannot be perfectly uncorrelated. In the general case, to find the resolution (PSF) of the covariance image, we need to calculate the spatial covariance of the patterns, which is the subject of Part 2b, below.

Part 2b: Calculate pattern-pattern covariance

To calculate the spatial covariance of the projected patterns, we first consider the pattern formation model. In our experiments, for example, we use a DMD to create random patterns at the sample plane. Assuming that the projected DMD pattern is sparse enough to avoid interference cross-terms, we can express our pattern under the incoherent model as

$$p_\ell(\mathbf{r}) = \iint t_\ell(\mathbf{r}') h_{\text{illu}}(\mathbf{r} - \mathbf{r}') d^2 \mathbf{r}', \quad (3.8)$$

where $t_\ell(\mathbf{r})$ is the ℓ -th pattern on the DMD. With this model, the pattern-pattern covariance is

$$\begin{aligned} \langle \Delta p_\ell(\mathbf{r}) \Delta p_\ell(\mathbf{r}') \rangle_\ell &= \iint \iint \langle \Delta t_\ell(\mathbf{r}_1) \Delta t_\ell(\mathbf{r}_2) \rangle_\ell h_{\text{illu}}(\mathbf{r} - \mathbf{r}_1) h_{\text{illu}}(\mathbf{r}' - \mathbf{r}_2) d^2 \mathbf{r}_1 d^2 \mathbf{r}_2 \\ &= \iint \iint \gamma_t \langle \Delta t_\ell^2(\mathbf{r}_1) \rangle_\ell \delta(\mathbf{r}_1 - \mathbf{r}_2) h_{\text{illu}}(\mathbf{r} - \mathbf{r}_1) h_{\text{illu}}(\mathbf{r}' - \mathbf{r}_2) d^2 \mathbf{r}_1 d^2 \mathbf{r}_2 \\ &\approx \alpha_t \iint h_{\text{illu}}(\mathbf{r} - \mathbf{r}_1) h_{\text{illu}}(\mathbf{r}' - \mathbf{r}_1) d^2 \mathbf{r}_1 = \alpha_t (h_{\text{illu}} \star h_{\text{illu}})(\mathbf{r} - \mathbf{r}'), \end{aligned} \quad (3.9)$$

where we have used an assumption that the DMD pattern values at position \mathbf{r}_1 and \mathbf{r}_2 are perfectly uncorrelated:

$$\langle \Delta t_\ell(\mathbf{r}_1) \Delta t_\ell(\mathbf{r}_2) \rangle_\ell = \gamma_t \langle \Delta t_\ell^2(\mathbf{r}_1) \rangle_\ell \delta(\mathbf{r}_1 - \mathbf{r}_2) \approx \alpha_t \delta(\mathbf{r}_1 - \mathbf{r}_2), \quad (3.10)$$

with γ_t being a constant that maintains unit consistency. This assumption is valid because the effective DMD pixel size is small compared to the FWHM of the optical system and we can control $\Delta t_\ell(\mathbf{r})$ to create an uncorrelated pattern. In the experiment, each position of $t_\ell(\mathbf{r})$ is an independent and identically distributed random variable. When the number of patterns is large enough, the variance $\langle \Delta t_\ell^2(\mathbf{r}_1) \rangle_\ell$ approaches the same constant for all the positions. We can then combine γ_t and the variance into a single constant α_t .

Ideally, we can assume $h_{\text{illu}}(\mathbf{r}) \approx h_{\text{det}}(\mathbf{r})$ when $\lambda_{\text{illu}} \approx \lambda_{\text{det}}$, where λ_{det} is the wavelength of the fluorescent emission detection light, and theoretically calculate the pattern-pattern covariance. We can also estimate $h_{\text{illu}} \star h_{\text{illu}}(\mathbf{r})$ by numerically evaluating Eq. (3.9) using our estimated patterns, which accounts for possible aberrations in the illumination optics.

Part 2c: PSF deconvolution of the covariance image

The pattern-pattern covariance derived in Part 2b is related to the PSF of the pattern-intensity covariance calculated in Part 2a. Hence, we can plug the pattern-pattern covariance into Eq. (3.7) and write the covariance image as

$$I_{\text{cov}}(\mathbf{r}) = \langle \Delta p_\ell(\mathbf{r}) \Delta I_\ell(\mathbf{r}) \rangle_\ell = \iint \alpha_t o(\mathbf{r}') [(h_{\text{illu}} \star h_{\text{illu}}) \cdot h_{\text{det}}](\mathbf{r} - \mathbf{r}') d^2 \mathbf{r}'. \quad (3.11)$$

Importantly, the effective PSF for this correlation image is now $[(h_{\text{illu}} \star h_{\text{illu}}) \cdot h_{\text{det}}](\mathbf{r})$, and the corresponding effective OTF is $[|\tilde{h}_{\text{illu}}|^2 \otimes \tilde{h}_{\text{det}}](\mathbf{u})$. Since both $|\tilde{h}_{\text{illu}}|^2$ and \tilde{h}_{det} have approximately the same Fourier support as the widefield OTF, the convolution between them covers around $2\times$ the support of the widefield OTF, as in conventional SIM. Given the effective PSF, we implement a standard deconvolution to improve contrast at high spatial frequencies:

$$I_{\text{cov,dec}}(\mathbf{r}) = \mathcal{F}^{-1} \left\{ \frac{\tilde{I}_{\text{cov}}(\mathbf{u}) \cdot H(\mathbf{u})}{|H(\mathbf{u})|^2 + \xi} \right\}, \quad (3.12)$$

where $H(\mathbf{u}) = [|\tilde{h}_{\text{illu}}|^2 \otimes \tilde{h}_{\text{det}}](\mathbf{u})$ and ξ is a small regularization parameter.

Part 2d: Shading correction operation

When the number of images is not large enough to give uniform variance of the patterns at each pixel ($\langle \Delta t_\ell^2(\mathbf{r}') \rangle_\ell$ from Eq. (3.9)), low-frequency shading artifacts will occur. These can be seen in the deconvolved covariance image in Fig. 3.3. To resolve this, we can estimate and correct for the variance across the image using our previously estimated projected patterns. Since the projected pattern $p_\ell(\mathbf{r})$ is the blurred version of the pattern on the DMD, by ignoring the high-frequency component of the DMD pattern, we can approximate the variance of the DMD pattern by

$$\alpha_t(\mathbf{r}) = \gamma_t \langle \Delta t_\ell^2(\mathbf{r}) \rangle_\ell \approx \gamma_t \langle \Delta p_\ell^2(\mathbf{r}) \rangle_\ell. \quad (3.13)$$

We divide out the spatially-varying variance α_t in Eq. (3.11) from the deconvolved SIMS image,

$$I_{\text{SIMS}}(\mathbf{r}) = \frac{I_{\text{cov,dec}}(\mathbf{r}) \cdot \alpha_t(\mathbf{r})}{\alpha_t^2(\mathbf{r}) + \epsilon}, \quad (3.14)$$

where ϵ is a regularizer and $I_{\text{SIMS}}(\mathbf{r})$ is the output from our SIMS reconstruction (Part 2c). This result of this step is our final reconstruction of the high-resolution object function.

3.2.4 Parameter tuning

Our SIMS algorithm involves 4 regularizers: β , δ , ξ , and ϵ , described in Eq. (3.3), Eq. (3.6), Eq. (3.12), and Eq. (3.14), respectively. Each is decoupled from the others and acts similarly to a typical Tikhonov regularizer, so tuning may be done independently. Generally, we want the regularizers to be as small as possible, while still avoiding noise amplification.

The procedure to tune the regularization parameters heuristically is summarized as follows. First, we check if the widefield images are well-deconvolved by finding the smallest β to give the image with best resolution but without obvious noise amplification, then we move on to check the deconvolved covariance image by tuning the SIMS regularizer ξ and the smooth-edge filter regularizer δ using the same principle, and finally we check the final reconstruction by using the smallest shading correction regularizer ϵ with enough shading correction but without evident noise amplification.

3.3 Simulation results

3.3.1 Definition of resolution

Before introducing and comparing any SIM algorithms, we want to first define the resolution criterion considered in this paper. Resolution of a microscopic image is usually defined by measuring the minimal resolvable distance between two points. Consider a widefield image with detected wavelength λ and numerical aperture NA ; the Abbe resolution criterion is then

$0.5\lambda/NA$, the full width at half maximum (FWHM) of the widefield PSF. As two points get closer to each other, the contrast between them decreases. Hence, the Abbe resolution criterion can be thought of as setting the minimum acceptable contrast between two points at 0.01. We can therefore define the resolution of a microscope or a reconstruction algorithm by measuring the smallest resolvable fine features that have contrast between them of at least 0.01.

3.3.2 Comparison of algorithms

Given this definition of resolution, we quantify the resolution for various algorithms in Fig. 3.4. The Siemens star test target ($o(r, \theta) = 1 + \cos 40\theta$ in polar coordinates) has varying spatial frequencies along the radius. The resolution of different imaging methods is quantified by reading the minimal resolved period when the contrast reaches 0.01. The effective modulation transfer function (MTF) of each method is shown in Fig. 3.4b, measured as the contrast of the reconstructed Siemens star image at different radii.

Our simulations use a SIM dataset with random patterns, so that we may compare against the previously proposed reconstruction algorithms of blind SIM [51] and S-SOFI [55]. We create $N_{\text{img}} = 400$ speckle-illuminated images from shifted random patterns on the DMD, with shifts of 0.6 FWHM of the PSF across 20×20 steps in the x and y directions, respectively. In each pattern, only 10% of the DMD pixels are turned on. This noise-free situation allows us to compare the ideal achieved resolution for the different algorithms.

Figure 3.4a shows the widefield, deconvolved widefield, confocal, and deconvolved confocal images of the Siemens star, as compared to blind SIM [51], S-SOFI [55] and our algorithm. At the bottom, we show the measured effective MTF for each algorithm. In terms of visual effect, S-SOFI [55] gives the least artifacts.

Table 3.1: Achieved resolution for different algorithms

	Widefield	Widefield deconvolved	Confocal	Confocal deconvolved
Resolution [$\lambda/2NA$]	1.035	0.844	0.681	0.428
Enhancement	$1 \times$	$1.23 \times$	$1.52 \times$	$2.42 \times$
	Blind SIM	S-SOFI	PE-SIMS	PE-SIMS-PR
Resolution [$\lambda/2NA$]	0.563	0.619	0.551	0.517
Enhancement	$1.84 \times$	$1.67 \times$	$1.88 \times$	$2.00 \times$

To compare resolution, we use our definition of the minimal resolved separation when the contrast drops to 0.01 and summarize the results in Table 3.1. The enhancement metric gives the ratio resolution improvement over widefield imaging. S-SOFI resolves features down to $1.67 \times$ smaller than the widefield microscope, which is close to the claimed $1.6 \times$ in [55],

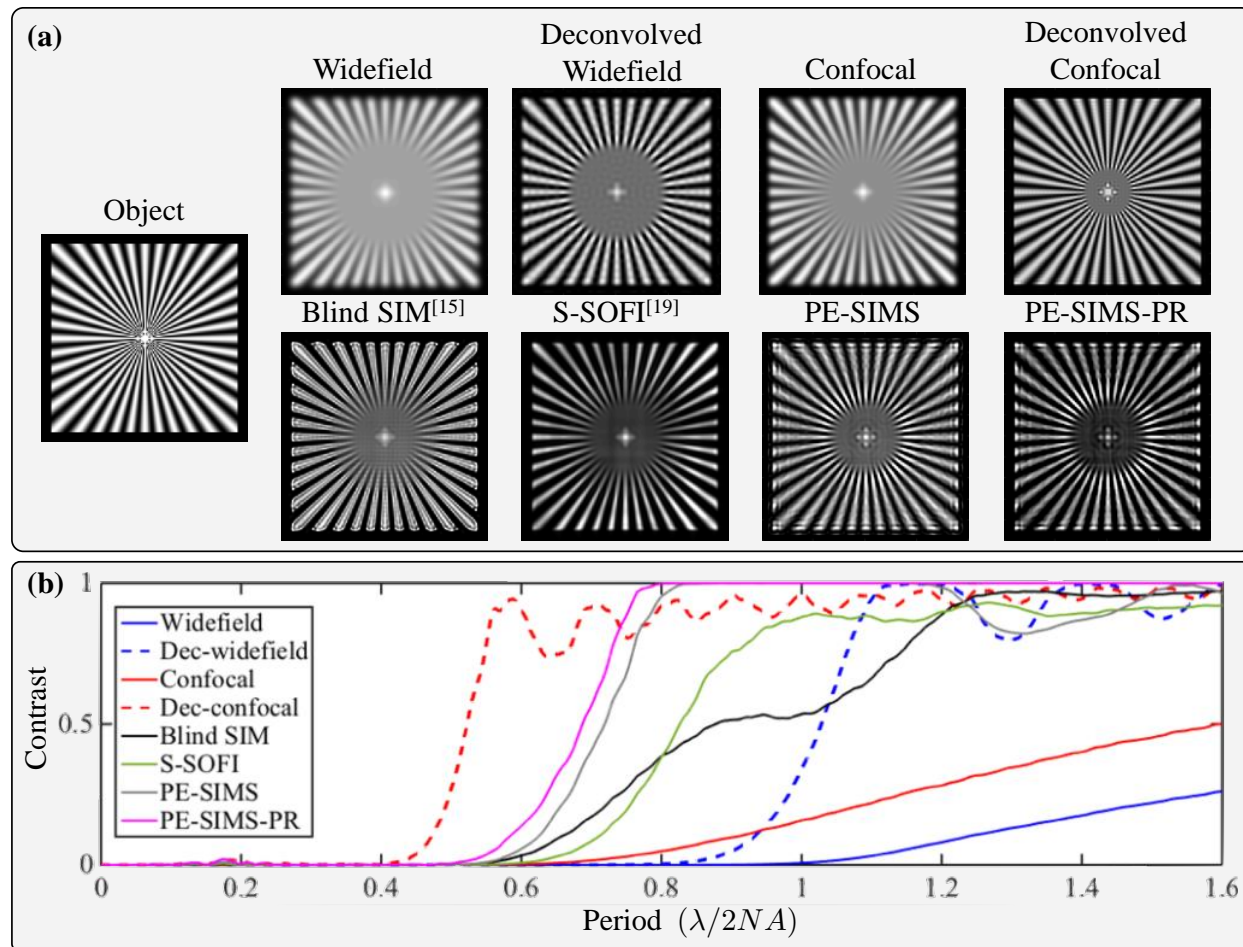


Figure 3.4: (a) Simulated reconstructions of a Siemens star target under a widefield microscope, deconvolved widefield, confocal microscope, deconvolved confocal, blind SIM [51], S-SOFI [55], our PE-SIMS and PE-SIMS-PR algorithms. (b) The effective modulation transfer function (MTF) of each method, given by the contrast of the reconstructed Siemens star image at different radii.

and Blind SIM achieves $1.84\times$ improvement but lower contrast for high-frequencies, which is consistent with [51]. Our PE-SIMS and PE-SIMS-PR (PE-SIMS with pixel reassignment algorithm [116–120] described in Sec. 3.6) algorithms give better resolution compared to other methods. We resolve features down to $1.84\times$ and $2\times$, respectively, close to the limit set by the deconvolved confocal image. Hence, our method performs the best of the blind algorithms.

Ideally, if we know all the patterns and our spatial modulation covers the full Fourier bandwidth of the objective, we could reconstruct out to $4NA/\lambda$ in Fourier space, achieving enhancement of $2.42\times$, as in the case of deconvolved confocal image or periodic SIM with known patterns. The blind algorithms, however, deal with an ill-posed problem (measure

N_{img} images and solve $N_{\text{img}} + 1$ images) that can only become well-posed through appropriate constraints. If the prior for these algorithms are not accurate enough, they may solve a different problem even if the problem becomes well-posed. This is why algorithms with different prior assumptions give different resolution performance for the same dataset, as we saw in Table 3.1.

3.4 Experimental results

Our experimental setup is shown in Fig. 3.1. A laser beam (Thorlabs, CPS532, 4.5 mW) is expanded to impinge onto a reflective DMD spatial light modulator (DLP[®]Discovery 4100, .7 XGA, 1024×768 pixels, pixel size 13.6 μm). The DMD generates a total of N_{img} random patterns (30% of DMD pixels turned on). These random illumination patterns are projected onto the object (with demagnification of 60×) through a $4f$ system composed of a 200 mm convex lens and a 60× objective lens with NA= 0.8 (Nikon CFI). The resulting fluorescent light is then collected with another $4f$ system formed by the same 60× objective and a 400 mm convex lens (magnification 120×). A dichroic mirror blocks the reflected illumination light (as in a typical epi-illumination setup). The images are taken with an sCMOS camera (PCO.edge 5.5, 2560×2160 pixels, pixel size 6.5 μm). Patterns are shifted on a 20×20 grid in the x and y directions with a step size of 0.6 FWHM of the PSF, while collecting images at each step. Our test object is carboxylate-modified red fluorescent beads (Excitation wavelength: 580 nm/Emission wavelength: 605 nm) having mean diameter of 210 nm (F8810, Life Technologies).

Reconstruction results are shown in Fig. 3.5, demonstrating improved resolution using our PE-SIMS algorithm, as compared to standard widefield or deconvolved widefield images. To quantitatively analyze the experimental results, we measure the resolved feature size of the reconstructed image and compare it to our theory. As shown in the cutline in Fig. 3.5, two fluorescent beads separated by 328 nm can clearly be resolved using our method, which are otherwise unresolvable in either widefield or deconvolved widefield images. The contrast of this two-Gaussian shape shows these two Gaussian are separated by $1.16 \times$ FWHM, so the FWHM of the reconstructed beads is around 283 nm, which is below the diffraction limit $\lambda/2NA = 371$ nm and closer to the size of the beads (210 nm in diameter).

Our algorithm can be used on other types of SIM datasets, as long as the pattern-pattern covariance gives a point-like function at the center. As an example, we tested our algorithm on a dataset from a previous method, Multispot SIM (MSIM) [45]. In MSIM, the patterns are a shifting grid of diffraction-limited spots. Since the previous MSIM implementation assumes known patterns, a calibration step captured an extra dataset with a uniform fluorescence sample in order to measure the patterns directly. Our algorithm ignores this calibration data, yet accurately reconstructs both the object and patterns (see Fig. 3.6). The MSIM result using the calibration data is shown for comparison. The sample is microtubules stained with Alexa Fluor 488 in a fixed cell observed under a TIRF 60× objective with $NA = 1.45$. Our PE-SIMS-PR reconstruction gives a similar result to the known-pattern

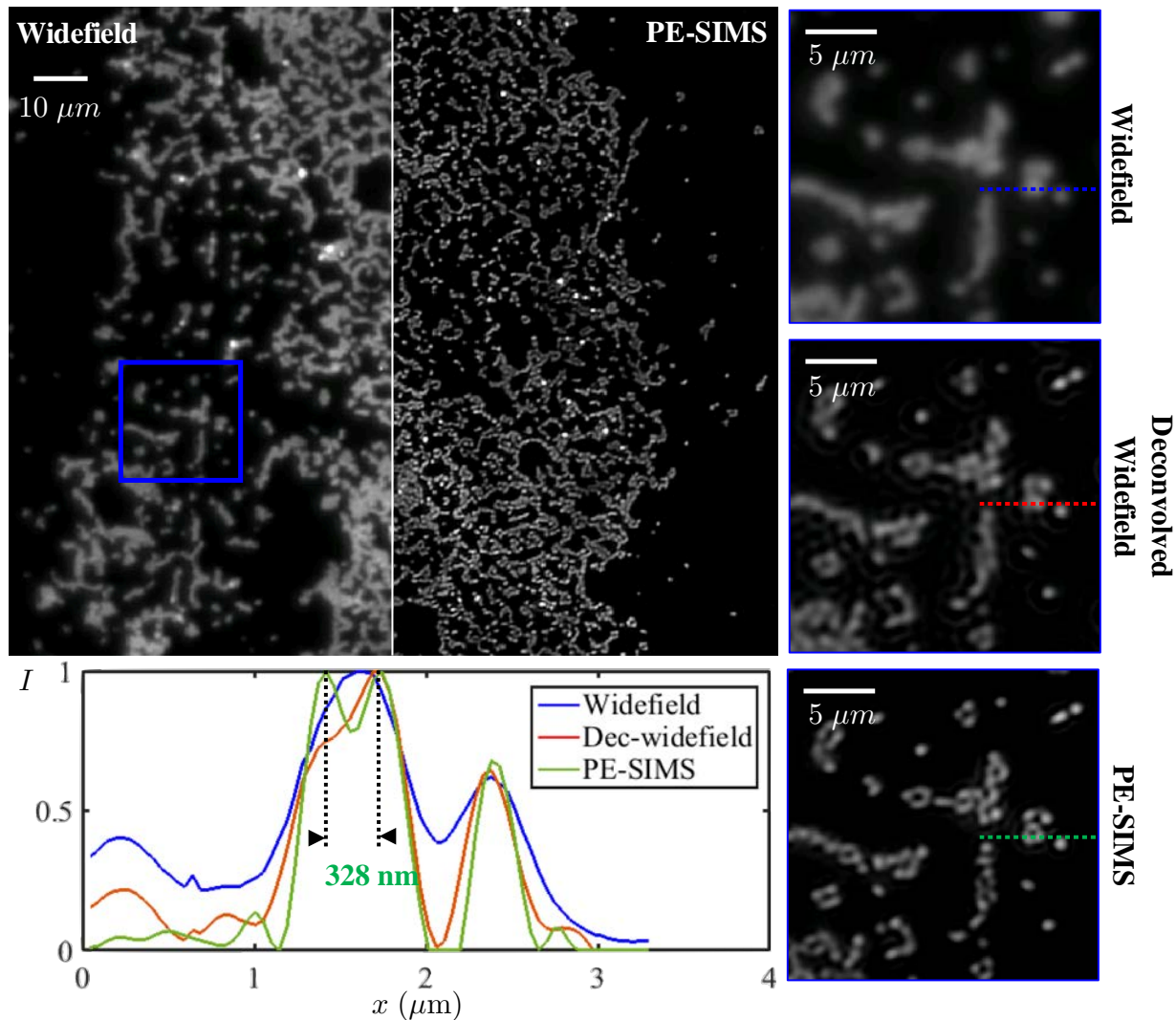


Figure 3.5: Reconstructions of red fluorescent beads (Ex:580 nm/Em:605 nm) from the experiment using random pattern illumination with 20×20 scanning step.

MSIM reconstruction.

3.5 Reduced acquisitions by multi-spot scanning

In this paper, we used 400 random speckle illumination patterns to reconstruct the image, far more than the 9-image requirement of conventional SIM [6]. This large number of images was required for high-quality reconstructions because the average and variance of the illumination patterns must be sufficiently flat in order to avoid shading variations. Recall that we want $\alpha_t(\mathbf{r}) \approx \gamma_t \langle \Delta p_\ell^2(\mathbf{r}) \rangle_\ell$ in Eq. (3.13) to be close to a constant, which suggests that the variance of

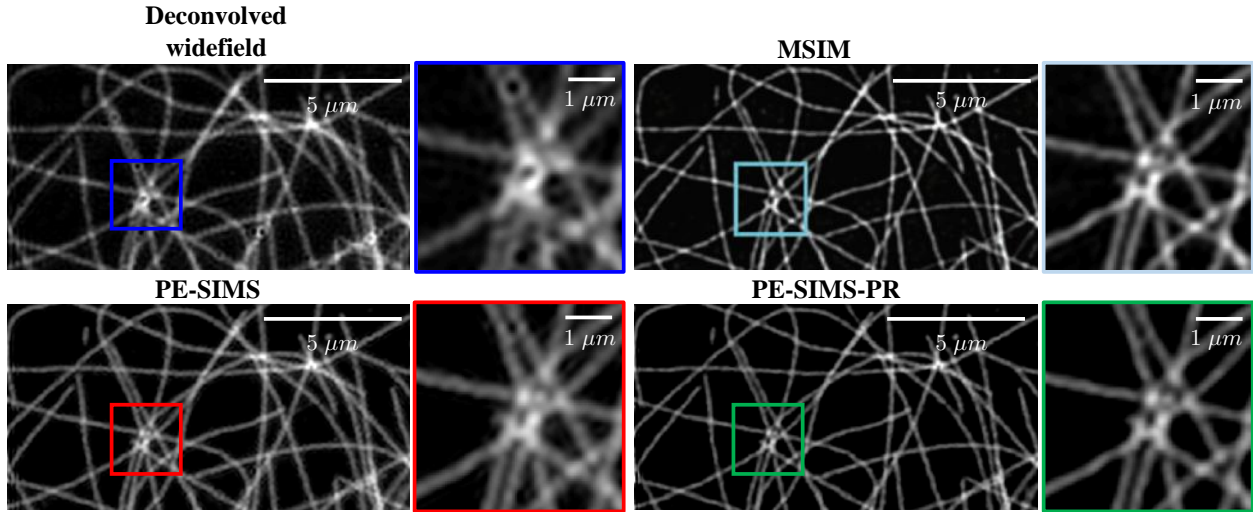


Figure 3.6: Comparison of our algorithm on dataset from Multispot SIM (MSIM) which uses with $N_{\text{img}} = 224$ scanned multi-spot patterns from [45]. We show the deconvolved widefield image and the reconstructions using MSIM with known patterns, as well as our blind PE-SIMS algorithm with and without pixel reassignment.

the random patterns is constant. When the number of images N_{img} goes down, this statistical assumption is not true any more. We use a shading correction algorithm (Sec. 3.2.3) to fix this problem by estimating the nonuniform variance, but it is still only an estimate. Hence, when the degree of variance nonuniformity increases (as the number of images decreases), the shading correction algorithm incurs errors.

Figure 3.7 shows simulations demonstrating the effect of reducing the number of images. We use the same random pattern as in Sec. 3.2.3 and shift by step sizes of 0.6 FWHM of the PSF. As we decrease the number of images from 400 to 36, the reconstruction becomes worse, due to shading errors. The shading map, $\alpha_t(\mathbf{r})o(\mathbf{r})$, is shown in the bottom row of Fig. 3.7. We can see the artifacts happen at the region where the $\alpha_t(\mathbf{r})$ is dim and changing. Without knowing the patterns a priori it is not possible to fully correct these shading effects.

Since we know that the artifacts that appear with too few images are due to a non-uniform $\alpha_t(\mathbf{r})$, we can attempt to design patterns that will be uniform with a minimal number of images. We would like $\langle \Delta p_\ell^2(\mathbf{r}) \rangle_\ell$ to give a uniform map. Consider the contribution from a single pattern; $\Delta p_\ell^2(\mathbf{r})$ is similar to the original pattern but with sharper bright spots. The ensemble average over ℓ sums up all these bright spots after shifting the pattern around. For a shifted random pattern, we must capture many images in order for the summation of the bright spots to give a uniform map. One efficient way to get a sum of bright spots to become a uniform map is to use a periodic multi-spot pattern (see Fig. 3.7) [45, 46, 48]. The period of this multi-spot pattern is designed to be 6 shifting step sizes. Thus, we can use 6×6 scanning steps to give a uniform shading map $\alpha_t(\mathbf{r})$. The reconstruction is also shown in Fig. 3.7 to be almost as good as the one illuminated with 400 shifted random patterns.

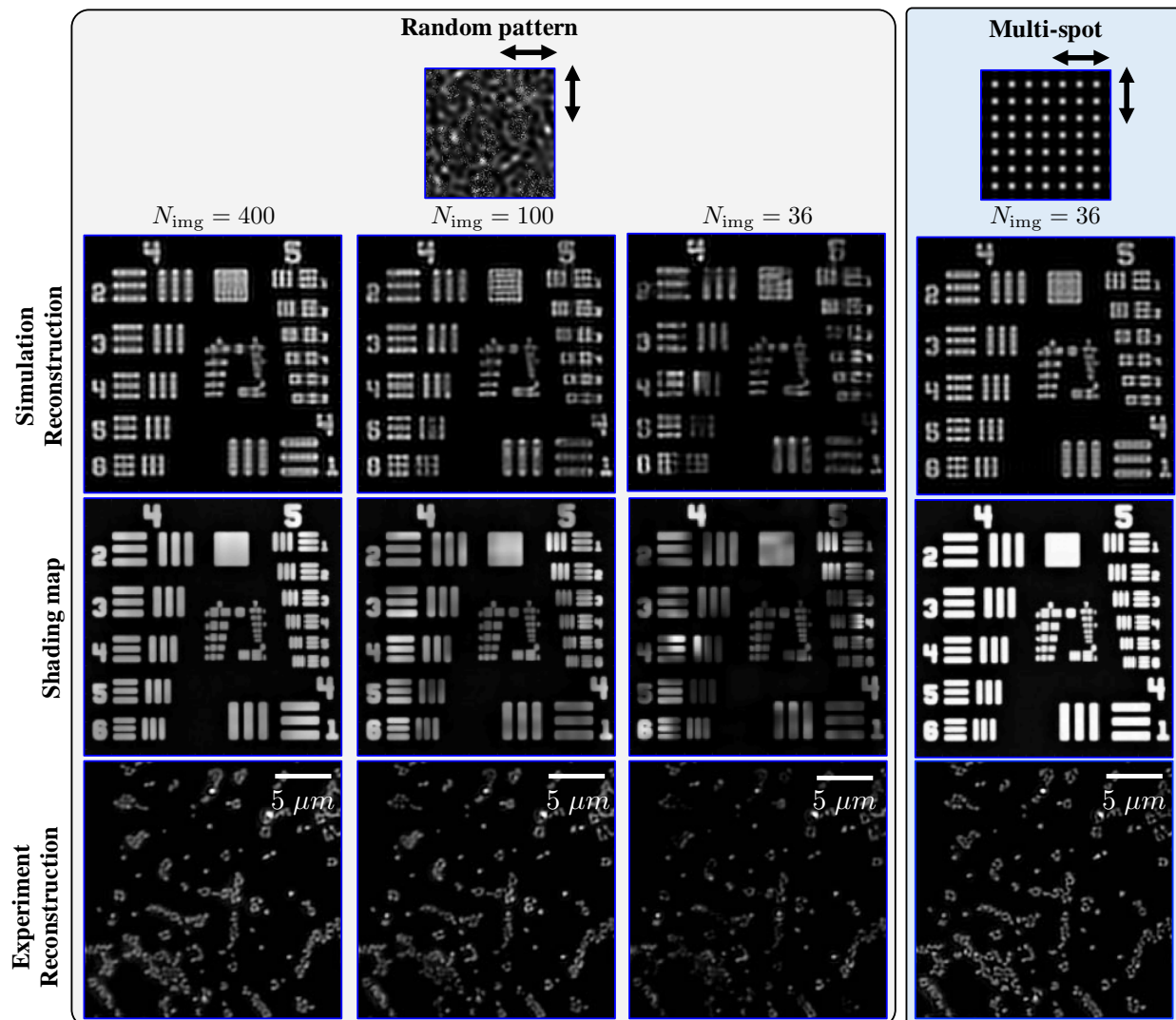


Figure 3.7: Results with simulated and experimental (fluorescent beads) datasets comparing random speckle and multi-spot illumination patterns. (middle row) Shading maps overlaid on the object. Decreasing the number of random patterns results in shading artifacts in the reconstruction. The random patterns are scanned in 20×20 , 10×10 , and 6×6 steps with the same step size of 0.6 FWHM of the PSF, while the multi-spot pattern is scanned with 6×6 steps.

Experimentally, we see similar trends in image reconstruction quality for different illumination strategies (see the bottom row of Fig. 3.7). Results from random pattern illumination of fluorescent beads with $N_{\text{img}} = 400$ and multi-spot illumination with $N_{\text{img}} = 36$ give very similar results, and shading artifacts become prominent as the number of patterns is reduced. Note that we use the same algorithm for both the random and multi-spot illuminated datasets

because the PSFs of the pattern-intensity covariance images $I_{\text{cov}}(\mathbf{r})$ for both cases are the same.

To show that the PSF for the pattern-intensity covariance image with random and multi-spot illumination are the same, we must derive the pattern-pattern covariance $\langle \Delta p_\ell(\mathbf{r}) \Delta p_\ell(\mathbf{r}') \rangle_\ell$ as we did in Part 2b in Sec. 3.2.3. To calculate the pattern-pattern covariance, we need to calculate the covariance of the patterns on the DMD $\langle \Delta t_\ell(\mathbf{r}) \Delta t_\ell(\mathbf{r}') \rangle_\ell$ and plug it into Eq. (3.9) to get pattern-pattern covariance $\langle \Delta p_\ell(\mathbf{r}) \Delta p_\ell(\mathbf{r}') \rangle_\ell$. For the multi-spot case, we can express the pattern on the DMD and its zero-mean pattern as

$$\begin{aligned} t_\ell(\mathbf{r}) &= \Lambda^2 \sum_{m,n} \delta(\mathbf{r} - \mathbf{r}_{mn} - \mathbf{r}_\ell) + t_0 \\ \Delta t_\ell(\mathbf{r}) &\approx \Lambda^2 \sum_{m,n} \delta(\mathbf{r} - \mathbf{r}_{mn} - \mathbf{r}_\ell), \end{aligned} \quad (3.15)$$

where $\mathbf{r}_{mn} = (m\Lambda, n\Lambda)$, m and n are integers, and Λ is the period of the pattern. Then, we can calculate the covariance of the pattern on the DMD as

$$\begin{aligned} \langle \Delta t_\ell(\mathbf{r}_1) \Delta t_\ell(\mathbf{r}_2) \rangle_\ell &= \iint \Delta t(\mathbf{r}_1 - \mathbf{r}_\ell) \Delta t(\mathbf{r}_2 - \mathbf{r}_\ell) d^2 \mathbf{r}_\ell \\ &= \Lambda^4 \sum_{m,n} \delta(\mathbf{r}_1 - \mathbf{r}_2 - \mathbf{r}_{mn}) \star \sum_{m,n} \delta(\mathbf{r}_1 - \mathbf{r}_2 - \mathbf{r}_{mn}) \\ &\approx \Lambda^4 \eta \sum_{m,n} \delta(\mathbf{r}_1 - \mathbf{r}_2 - \mathbf{r}_{mn}), \end{aligned} \quad (3.16)$$

where η is a constant that enforces unit consistency. Plugging this into Eq. (3.9), we can then calculate the pattern-pattern covariance as

$$\langle \Delta p_\ell(\mathbf{r}) \Delta p_\ell(\mathbf{r}') \rangle_\ell = (h_{\text{illu}} \star h_{\text{illu}})(\mathbf{r} - \mathbf{r}') \otimes \Lambda^4 \eta \sum_{m,n} \delta(\mathbf{r} - \mathbf{r}' - \mathbf{r}_{mn}). \quad (3.17)$$

Although the pattern-pattern covariance is only a replica of the $(h_{\text{illu}} \star h_{\text{illu}})(\mathbf{r})$, the PSF of the covariance image, $I_{\text{cov}}(\mathbf{r})$, only depends on the multiplication of $h_{\text{det}}(\mathbf{r})$ and $(h_{\text{illu}} \star h_{\text{illu}})(\mathbf{r}) \otimes \Lambda^4 \eta \sum_{m,n} \delta(\mathbf{r} - \mathbf{r}_{mn})$ as Eq. (3.7) derived. If the period of the multi-spot pattern is large compared to $(h_{\text{illu}} \star h_{\text{illu}})(\mathbf{r})$, we can still have our PSF as $[(h_{\text{illu}} \star h_{\text{illu}}) \cdot h_{\text{det}}](\mathbf{r})$, which is the same as the case of random pattern illumination.

3.6 Pixel reassignment for better SNR

In this section, we first discuss the similarity between SIMS and confocal microscopy. This leads to an extension of our method that incorporates the pixel reassignment procedure proposed in [116–120]. In computing the covariance of the shifted pattern $p_\ell(\mathbf{r} - \mathbf{r}_s)$ and the intensity $I_\ell(\mathbf{r})$, there is still some information of the object leftover. Pixel reassignment helps incorporate it in a straightforward fashion, giving better SNR in the final reconstruction.

In Sec. 3.2.3 of our SIMS procedure, we first calculate the covariance image $I_{\text{cov}}(\mathbf{r})$. The PSF of this covariance image is determined by imposing our statistical prior on the pattern-pattern covariance $\langle \Delta p_\ell(\mathbf{r}) \Delta p_\ell(\mathbf{r}') \rangle_\ell$. The effect is similar to the illumination PSF of confocal microscopy [44]. Looking at Eq. (3.11), our covariance image with PSF of $[(h_{\text{illu}} \star h_{\text{illu}}) \cdot h_{\text{det}}](\mathbf{r})$ is the same as a confocal image taken with illumination PSF, $(h_{\text{illu}} \star h_{\text{illu}})(\mathbf{r})$, and detection PSF, $h_{\text{det}}(\mathbf{r})$.

From the same SIM dataset, we can further use the shifted patterns $p_\ell(\mathbf{r} - \mathbf{r}_s)$ and correlate them with the intensity $I_\ell(\mathbf{r})$ to compute a series of shifted covariance images

$$\begin{aligned} I_{\text{cov}}^s(\mathbf{r}, \mathbf{r}_s) &= \langle \Delta p_\ell(\mathbf{r} - \mathbf{r}_s) \Delta I_\ell(\mathbf{r}) \rangle_\ell = \iint o(\mathbf{r}') \langle \Delta p_\ell(\mathbf{r} - \mathbf{r}_s) \Delta p_\ell(\mathbf{r}') \rangle_\ell h(\mathbf{r} - \mathbf{r}') d^2 \mathbf{r}' \\ &= \iint \alpha_t o(\mathbf{r}') (h_{\text{illu}} \star h_{\text{illu}})(\mathbf{r} - \mathbf{r}_s - \mathbf{r}') h_{\text{det}}(\mathbf{r} - \mathbf{r}') d^2 \mathbf{r}'. \end{aligned} \quad (3.18)$$

The PSF of the shifted covariance image $I_{\text{cov}}^s(\mathbf{r})$ is the product of $(h_{\text{illu}} \star h_{\text{illu}})(\mathbf{r} - \mathbf{r}_s)$ and $h_{\text{det}}(\mathbf{r})$, whose center is approximately at $\mathbf{r}_s/2$. This image is the same as the image taken under a confocal microscope with a shifted pinhole. This implies by shifting around the patterns and correlating with the intensity, we get the equivalent of many 2D confocal images taken with the pinhole at different positions. This is the same dataset as would be described in the imaging scanning microscope, where the single-pixel camera and pinhole is replaced with a CCD in the confocal system [118, 119]. Though these images are not centered, they still contain the information of the same object. Pixel reassignment was proposed in [117–120] as a way to incorporate this 4D information to get a 2D image with better SNR.

Since the 2D images from \mathbf{r}_s -shifted patterns are approximately $\mathbf{r}_s/2$ -shifted versions of the one at $\mathbf{r}_s = \mathbf{0}$, we can shift the information back to the center region and sum up all these images to enhance the SNR and form a pixel-reassigned (PR) image as

$$\begin{aligned} I_{\text{PR}}(\mathbf{r}) &= \iint I_{\text{cov}}^s \left(\mathbf{r} + \frac{\mathbf{r}_s}{2}, \mathbf{r}_s \right) d^2 \mathbf{r}_s \\ &= \iint \alpha_t o(\mathbf{r}') \left[\iint (h_{\text{illu}} \star h_{\text{illu}}) \left(\mathbf{r} - \frac{\mathbf{r}_s}{2} - \mathbf{r}' \right) h_{\text{det}} \left(\mathbf{r} + \frac{\mathbf{r}_s}{2} - \mathbf{r}' \right) d^2 \mathbf{r}_s \right] d^2 \mathbf{r}' \\ &= \iint \alpha_t o(\mathbf{r}') [(h_{\text{illu}} \star h_{\text{illu}}) \otimes h_{\text{det}}](2(\mathbf{r} - \mathbf{r}')) d^2 \mathbf{r}' \end{aligned} \quad (3.19)$$

This synthesized image using pixel reassignment gives a PSF of $[(h_{\text{illu}} \star h_{\text{illu}}) \otimes h_{\text{det}}](2\mathbf{r})$. Figure 3.8(a) shows the comparison between the SIMS PSF, $[(h_{\text{illu}} \star h_{\text{illu}}) \cdot h_{\text{det}}](\mathbf{r})$, and the PSF of SIMS with pixel reassignment, $[(h_{\text{illu}} \star h_{\text{illu}}) \otimes h_{\text{det}}](2\mathbf{r})$ both in the real space and the Fourier space (assuming $h_{\text{illu}} \approx h_{\text{det}}$). In the real space, the PSF after doing pixel reassignment looks fatter than the one without pixel reassignment. However, the OTF of the one with pixel reassignment has larger value in the high-frequency region, where the noise severely degrade the image resolution. Thus, we get better SNR by summing up all the information we have and have a OTF that better deals with noise at high-frequency region. Since we know the PSF, $[(h_{\text{illu}} \star h_{\text{illu}}) \otimes h_{\text{det}}](2\mathbf{r})$, and the shading map, $\alpha_t(\mathbf{r})$, of

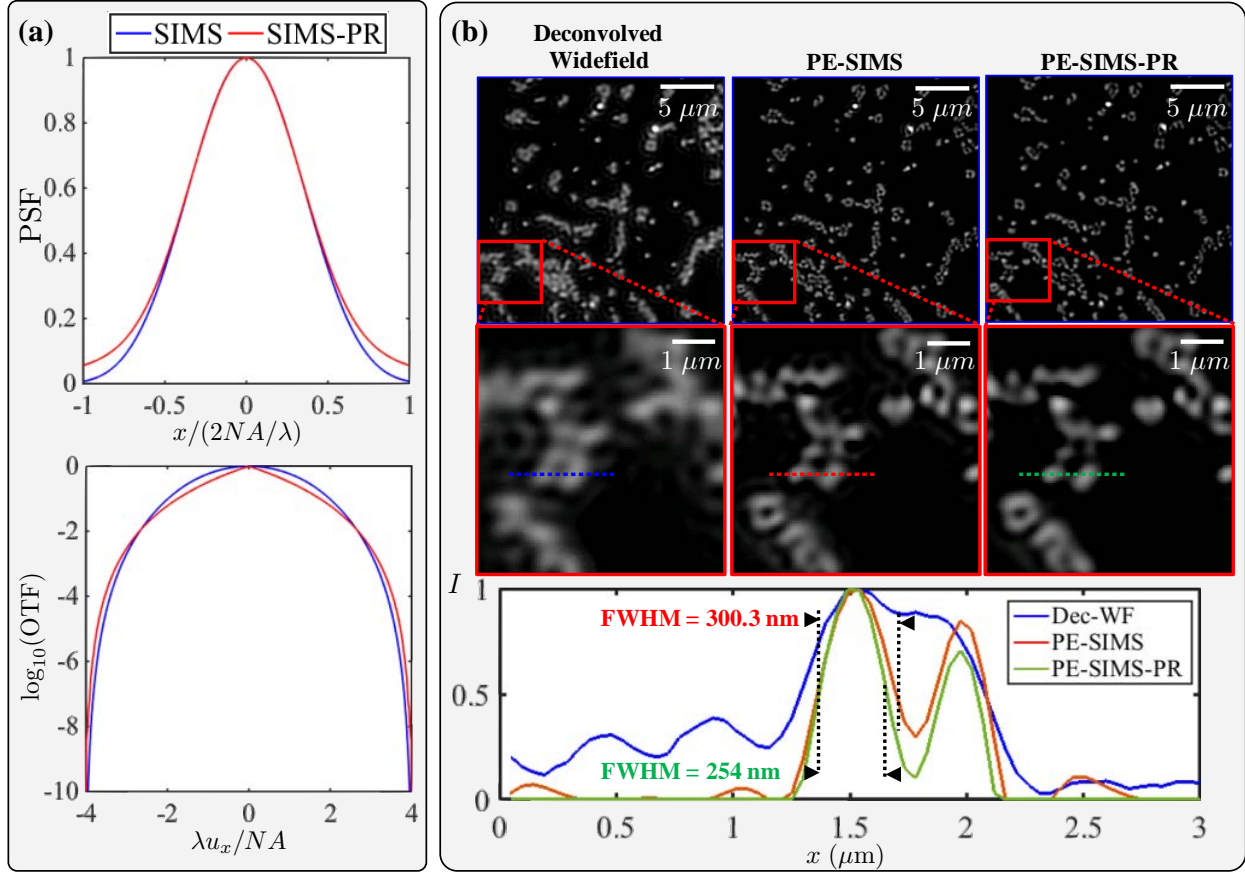


Figure 3.8: (a) Comparison of the PSF and OTF for SIMS and SIMS with pixel reassignment (PR). (b) Comparisons of the deconvolved widefield image and the reconstructions of the 6×6 multi-spot scanned fluorescent beads with and without pixel reassignment.

this pixel-reassigned image $I_{PR}(\mathbf{r})$, we can again apply the deconvolution and the shading correction operation described in Sec. 3.2.3 to get a PE-SIMS-PR reconstruction.

Figure 3.8(b) compares the reconstruction result of fluorescent beads using 6×6 multi-spot illumination with and without applying pixel reassignment algorithm. Pixel reassignment results in sharper contrast when two beads are close to each other and helps clean up some background deconvolution errors. A cut-line plot of the fluorescent beads in Fig. 3.8(b) shows that the FWHM of the reconstructed bead from SIMS (300.3 nm) is larger than for SIMS-PR with pixel reassignment (254 nm), giving better resolution.

3.7 Extension to 3D imaging

So far, we have seen our proposed algorithm effectively reconstructs 2D super-resolution fluorescent images without knowing the structured patterns, which relaxes the hardware

calibration requirement for SIM. Furthermore, the blind SIM idea is also applicable in 3D imaging as we can see in the demonstrations of both 3D-S-SOFI [128] and 3D blind SIM [56]. However, the previous 3D blind SIM imaging formulation is not accurate because modeling axial sample scanning requires special z -coordinate treatment mathematically. The mathematical structure of 2D SIM cannot be directly generalize to 3D by simply adding one more dimension. Here, we adopt more rigorous 3D imaging formulation as in [129] and propose a 3D extension of our 2D blind SIM algorithm to reconstruct 3D SIM images with unknown patterns. Finally, we apply this algorithm to experimental data with multi-spot illumination patterns to demonstrate 3D resolution enhancement without knowing the illumination patterns *a priori*.

3.7.1 Theory and method

Consider the experimental setup as in the 2D case shown in Fig. 3.1, except the sample here is mounted on a axially translating motion stage. A deformable mirror device (DMD) is used to project patterns onto the object through an objective lens, while a motion stage scans the sample in axial direction to collect an intensity stack. The ℓ -th captured intensity stack, $I_\ell(\mathbf{r}_\perp, z)$ (where $\mathbf{r}_\perp = (x, y)$ denotes lateral coordinates), can be represented as the product of the object's fluorescence distribution, $o(\mathbf{r}_\perp, z)$, and the illumination pattern, $p_\ell(\mathbf{r}_\perp, z)$, followed by a convolution with the system's detection-side incoherent 3D point spread function (PSF), $h_{\text{det}}(\mathbf{r}_\perp, z)$:

$$I_\ell(\mathbf{r}_\perp, z) = \iiint o(\mathbf{r}'_\perp, z') p_\ell(\mathbf{r}'_\perp, z - z') h_{\text{det}}(\mathbf{r}_\perp - \mathbf{r}'_\perp, z - z') d^2\mathbf{r}'_\perp dz'. \quad (3.20)$$

Note that the z -coordinate of the pattern in this model is not referenced to the object, but rather the PSF. This is because the pattern stays static relative to the objective lens while we scan the object axially. With this forward model, we can describe our self-calibration algorithm, 3D PE-SIMS, in two parts. The first part is Pattern Estimation (PE), in which an optimization procedure estimates the 3D illumination patterns. The second part is 3D SIM with a Statistical prior (SIMS), in which the super-resolved image is reconstructed from the estimated patterns and the measured intensity stacks.

3.7.1.1 Pattern Estimation

Part 1a: Approximate the widefield image

The captured intensity stack mixes information from the object and the illumination pattern. To recover the illumination patterns, which are diffraction-limited, we first make a low-resolution (diffraction-limited) estimate of the object by taking a mean of the intensity stacks:

$$I_{\text{avg}}(\mathbf{r}_\perp, z) = \langle I_\ell(\mathbf{r}_\perp, z) \rangle_\ell \approx p_0 o(\mathbf{r}_\perp, z) \otimes h_{\text{det}}(\mathbf{r}_\perp, z), \quad (3.21)$$

where $\langle \cdot \rangle_\ell$ denotes a mean operation with respect to ℓ . The mean of the illumination patterns should be close to a constant across the whole volume $\langle p_\ell(\mathbf{r}'_\perp, z - z') \rangle_\ell \approx p_0$, as in blind SIM [51].

Part 1b: Deconvolve the widefield image

The best estimate of the object before doing the SIMS reconstruction is obtained by a 3D deconvolution procedure on the mean intensity stack. This gives a low-resolution object estimate:

$$o_{\text{est}}(\mathbf{r}_\perp, z) = \mathcal{F}^{-1} \left\{ \frac{\tilde{I}_{\text{avg}}(\mathbf{u}_\perp, u_z) \cdot \tilde{h}_{\text{det}}(\mathbf{u}_\perp, u_z)}{|\tilde{h}_{\text{det}}(\mathbf{u}_\perp, u_z)|^2 + \beta} \right\}, \quad (3.22)$$

where $\mathbf{u}_\perp = (u_x, u_y)$ represents the lateral spatial frequency coordinates and β is the Tikhonov regularization constant.

Part 1c: Estimate patterns

The low-resolution object estimate $o_{\text{est}}(\mathbf{r}_\perp, z)$ is then used to estimate the patterns. In order to implement an efficient pattern estimation algorithm, we turn (3.20) into a convolution-based operation as

$$\begin{aligned} I_\ell(\mathbf{r}_\perp, z) &\approx \sum_{u'_z} \Delta u'_z \iint o(\mathbf{r}'_\perp, z') p_\ell(\mathbf{r}'_\perp, u'_z) h_{\text{det}, u'_z}(\mathbf{r}_\perp - \mathbf{r}'_\perp, z - z') d^2 \mathbf{r}'_\perp \\ &= \sum_{u'_z} \Delta u'_z [(o \cdot p_{\ell, u'_z}) \otimes h_{\text{det}, u'_z}](\mathbf{r}_\perp, z), \end{aligned} \quad (3.23)$$

where we have used

$$p_\ell(\mathbf{r}'_\perp, z - z') \approx \sum_{u'_z} \Delta u'_z p_\ell(\mathbf{r}'_\perp, u'_z) e^{i2\pi u'_z(z - z')}, \quad (3.24)$$

and defined $h_{\text{det}, u'_z}(\mathbf{r}_\perp, z) = h_{\text{det}}(\mathbf{r}_\perp, z) e^{i2\pi u'_z z}$ and $p_{\ell, u'_z} = p_\ell(\mathbf{r}_\perp, u'_z)$. A key feature of our algorithm is that we use this expression and a Fourier support constraint on the illumination patterns to estimate the 3D patterns via an optimization procedure

$$\min_{p_{\ell, u_z}} f(p_{\ell, u_z}) = \sum_{\mathbf{r}_\perp, z} \left| I_\ell(\mathbf{r}_\perp, z) - \sum_{u_z} \Delta u_z [(o \cdot p_{\ell, u_z}) \otimes h_{\text{det}, u_z}](\mathbf{r}_\perp, z) \right|^2 + \mathbb{I}_{\mathcal{C}}(p_{\ell, u_z}), \quad (3.25)$$

where $\mathbb{I}_{\mathcal{C}}(p_{\ell, u_z}) = \begin{cases} 0, & p_{\ell, u_z} \in \mathcal{C} \\ +\infty, & p_{\ell, u_z} \notin \mathcal{C} \end{cases}$, $\mathcal{C} = \left\{ p_{\ell, u_z} \mid \tilde{p}_\ell(\mathbf{u}_\perp, u_z) = 0, \text{ if } \tilde{h}_{\text{illu}}(\mathbf{u}_\perp, u_z) = 0 \right\}$,

where $\tilde{h}_{\text{illu}}(\mathbf{u}_\perp, u_z)$ represents the illumination-side incoherent 3D OTF. This optimization problem is solved by using a projected Nesterov's gradient descent method [123].

3.7.1.2 3D SIM with a Statistical prior

The second part of our algorithm uses the structured illumination dataset and the recovered illumination patterns to reconstruct the high-resolution object image.

Part 2a: Calculate pattern-intensity covariance

First, we calculate the pattern-intensity covariance as in 2D case [130]. Since we have estimated the patterns, we can retrieve the high-resolution object as

$$\begin{aligned}
 I_{\text{cov}}(\mathbf{r}_{\perp}, z) &\equiv \langle \Delta p_{\ell}(\mathbf{r}_{\perp}, 0) \Delta I_{\ell}(\mathbf{r}_{\perp}, z) \rangle_{\ell} \\
 &= \iiint o(\mathbf{r}'_{\perp}, z') \langle \Delta p_{\ell}(\mathbf{r}_{\perp}, 0) \Delta p_{\ell}(\mathbf{r}'_{\perp}, z - z') \rangle_{\ell} h_{\text{det}}(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z') d\mathbf{r}'_{\perp} dz' \\
 &= \iiint o(\mathbf{r}'_{\perp}, z') [(h_{\text{illu}} \star h_{\text{illu}}) \cdot h_{\text{det}}](\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z') d\mathbf{r}'_{\perp} dz',
 \end{aligned} \tag{3.26}$$

where we have used the pattern-pattern covariance $\langle \Delta p_{\ell}(\mathbf{r}_{\perp}, z) \Delta p_{\ell}(\mathbf{r}'_{\perp}, z') \rangle_{\ell} = (h_{\text{illu}} \star h_{\text{illu}})(\mathbf{r}_{\perp} - \mathbf{r}'_{\perp}, z - z')$. This can be considered as a 3D extension from the 2D correlation [130].

Part 2b: PSF deconvolution of covariance image

From 2a, the effective 3D PSF for the covariance image stack is $h_{\text{eff}}(\mathbf{r}_{\perp}, z) = [(h_{\text{illu}} \star h_{\text{illu}}) \cdot h_{\text{det}}](\mathbf{r}_{\perp}, z)$. Since we know the PSF h_{illu} and h_{det} , we can then conduct another deconvolution to get the final reconstruction:

$$I_{\text{cov,dec}}(\mathbf{r}_{\perp}, z) = \mathcal{F}^{-1} \left\{ \frac{\tilde{I}_{\text{cov}}(\mathbf{u}_{\perp}, u_z) \cdot \tilde{h}_{\text{eff}}(\mathbf{u}_{\perp}, u_z)}{|\tilde{h}_{\text{eff}}(\mathbf{u}_{\perp}, u_z)|^2 + \xi} A(\mathbf{u}_{\perp}, u_z) \right\}, \tag{3.27}$$

where ξ is another Tikhonov regularization constant and $A(\mathbf{u}_{\perp}, u_z)$ is the apodization filter (typically a triangular shape filter from the origin of Fourier space to the support of the effective OTF that approximates the extended OTF support).

3.7.2 Experimental results

Our experiment setup is geometrically the same as in the 2D case shown in Fig. 3.1, except that it is a high-NA version of the 2D setup and the object here is mounted on a translation stage for axial scanning. A 70 mW green laser beam at 532 nm is coupled into a single-mode fiber (Thorlabs, SM450) with an output power of 10 mW. The output beam is then collimated and impinge onto a reflective DMD spatial light modulator (DLP®Discovery 4100, .7" XGA, 1024×768 pixels, pixel size 13.6 μm). The patterns generated by the DMD are projected onto the object (with 125× demagnification) through a 4*f* system composed of a 250 mm convex lens (Thorlabs, AC508-250-A-ML) and a objective lens (Nikon, CFI Plan Apo Lambda 100X Oil, *NA* = 1.45). The resulting fluorescent light is then collected by the

same system with a dichroic mirror in the way to block the reflected illumination light. In the acquisition process, the object is scanned in the z -direction with a piezo-stage (Thorlabs, MAX311D) and the images are taken with an sCMOS camera (PCO.edge 5.5, 2560×2160 pixels, pixel size $6.5 \mu\text{m}$).

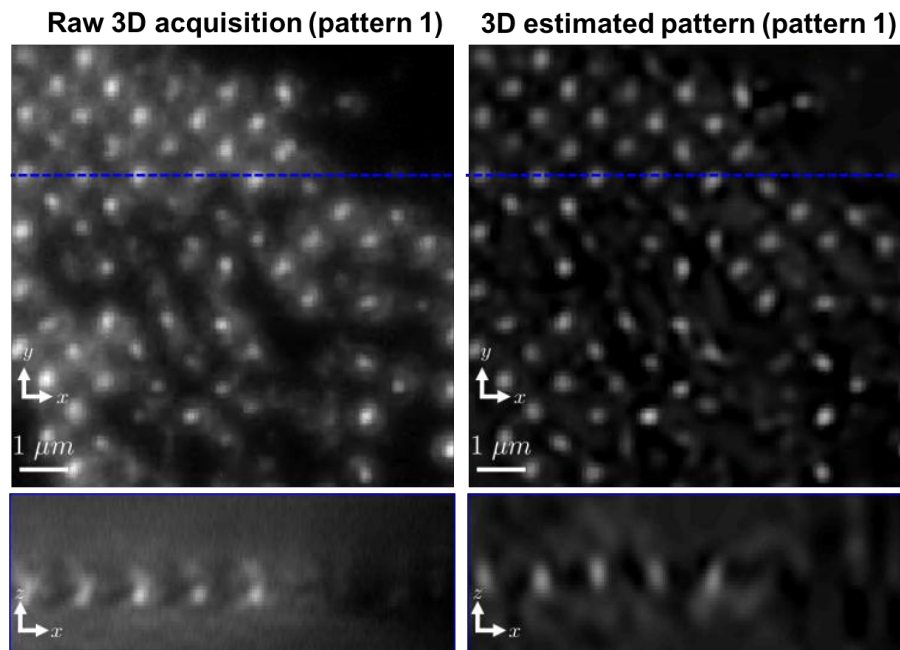


Figure 3.9: (Left) x - y and x - z cross-sections of the experimental raw data from fluorescent beads illuminated with a multispot pattern. (Right) The estimated 3D illumination pattern from our pattern estimation algorithm.

To first test out our algorithm and quantify the performance of it, we choose to image a test object, carboxylate-modified red fluorescent beads (Excitation wavelength: 580 nm/Emission wavelength: 605 nm) having mean diameter of 210 nm (F8810, Life Technologies). In this experiment, we generate a multispot pattern with period of 4.2 diffraction spot ($\lambda/2NA = 208.6$ nm) corresponding to 8 pixels on the DMD. This pattern is then shifted in a 8×8 grid with a step size of 0.525 diffraction spot and there are total 64 images taken for one z plane. This process is then repeated for 20 z planes. Figure 3.9 shows one of the x - y and x - z views of a 3D acquisition under one pattern illumination. It is obvious that the pattern information can be directly seen on the raw data. By plugging all the data into the first part of our algorithm, we can estimate the illumination pattern and the result is shown in the right hand side of Fig. 3.9. Compared with the raw data on the left, the estimated pattern has not only sharper illumination spots but also lower background signal.

With the estimated pattern, we then run the second part of our algorithm and give a 3D super-resolved reconstruction on the fluorescent beads. Figure 3.10 shows a comparison of our reconstruction and widefield image in one of the x - y and x - z views. Before the

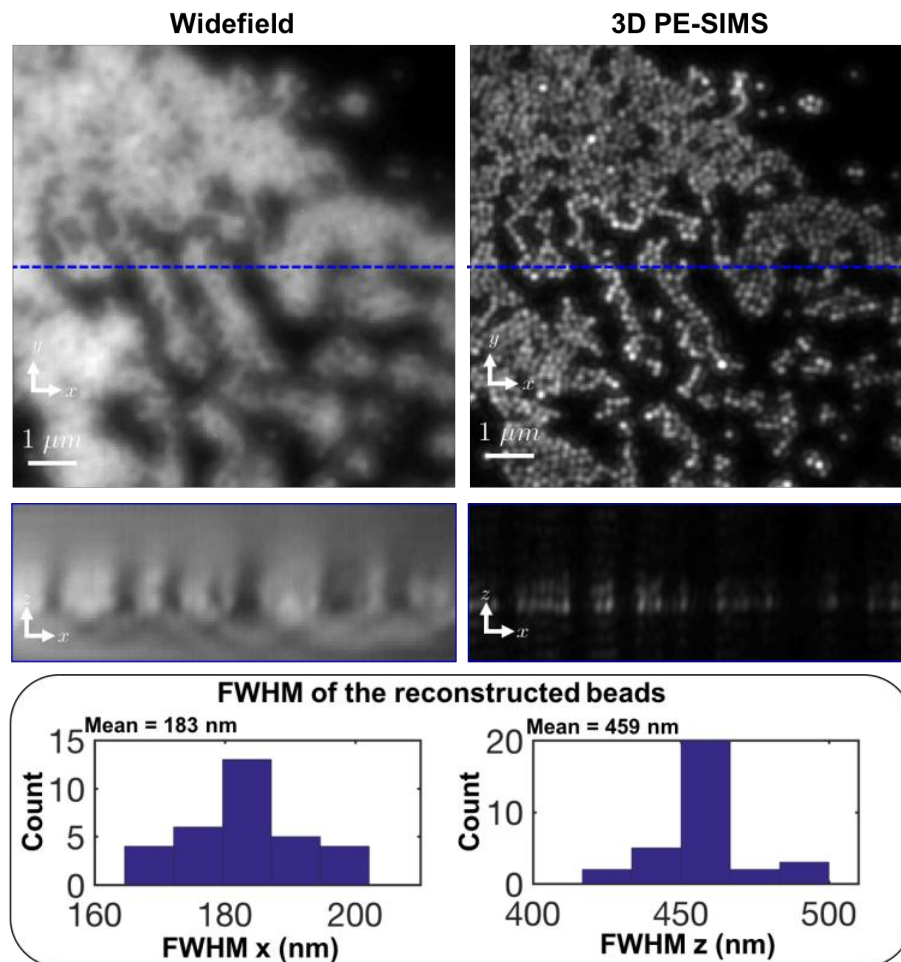


Figure 3.10: (Top) Comparison of widefield and 3D PE-SIMS reconstruction, showing x - y and x - z cross-sections for a sample with 200 nm red fluorescent beads. (Bottom) Histogram on the size (FWHM) of 32 reconstructed fluorescent beads.

reconstruction, the widefield image can not see each individual bead, while the reconstructed image can resolve almost every one of them.

To carefully quantify the resolution of our experiment, we randomly pick 32 fluorescent beads in the reconstruction and then calculate its full width at half maximum (FWHM) in both x and z direction. The bottom part of Fig. 3.10 shows the histogram and mean of the FWHM of the beads in x and z direction. However, this FWHM includes the effects from the size of the beads and size of the reconstructed PSF. We need to have the FWHM of a single bead to characterize the size of the reconstructed PSF. In the widefield image, the FWHM of an isolated fluorescent bead is 247 nm and a diffraction spot size (FWHM of widefield PSF) is 208 nm. We can then calculate the FWHM of a single 210 nm fluorescent bead as $\sqrt{247^2 - 208^2} = 133$ nm. With the FWHM of the fluorescent bead, we can then

characterize the FWHM of our reconstructed PSF as $\sqrt{183^2 - 133^2} = 126$ nm in x, y direction and $\sqrt{459^2 - 133^2} = 439$ nm in the z direction.

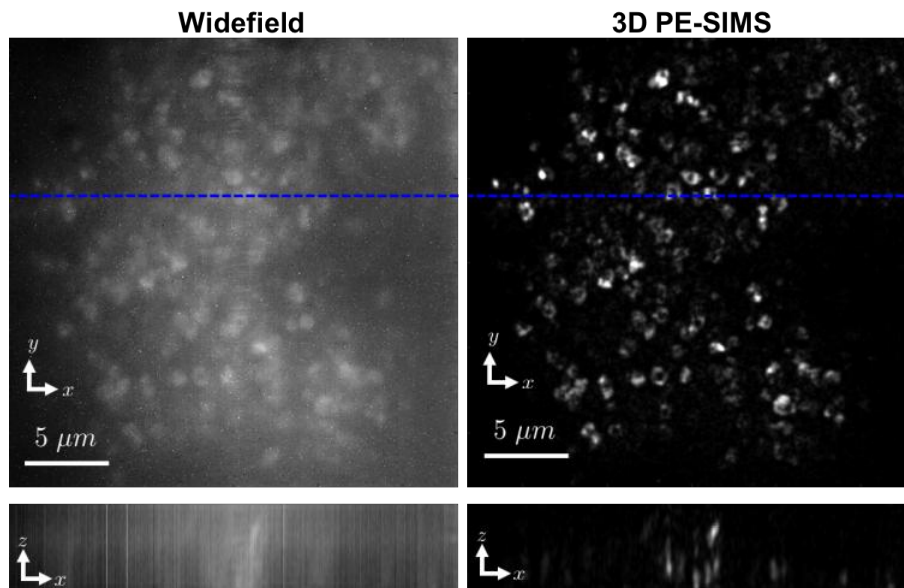


Figure 3.11: Comparison of widefield and 3D PE-SIMS reconstruction, showing x - y and x - z cross-sections for a 3D bacteria aggregation sample (*Candidatus Kuenenia stuttgartiensis*).

Next, we move on to image a more complicated biological sample (*Candidatus Kuenenia stuttgartiensis*) with the same setup and illumination patterns as our previous experiment. This sample is a 3D aggregation of bacteria with diameter of $1 \mu\text{m}$ [131]. The outer region of the bacteria is filled with Cy3-fluorescent-tagged 16s ribosomal RNA. Figure 3.11 shows a comparison between the widefield and 3D PE-SIMS reconstructed images of this bacteria in one of the x - y and x - z views. A clear resolution improvement in x and y and the depth sectioning ability are observed in the reconstructed images. To even verify this reconstruction, we can compare the reconstructed bacteria with the one observed under an electron microscope in [131] and find out that the size of both images match nicely.

3.8 Conclusion

We have proposed a robust algorithm that can give $2\times$ resolution improvement compared to widefield fluorescence imaging using 2D and 3D SIM dataset without knowing the imposed patterns. Our algorithm first estimates each illumination pattern from a low-resolution approximate object and measured intensities by solving a constrained convex optimization problem. We then synthesize a high-resolution image by calculating the covariance between the estimated patterns and the measured intensity images, followed by a deconvolution to get to the final reconstruction. For 2D algorithm, we quantified the limits on resolution of

our algorithm by the reconstructed contrast of a simulated Siemens star target. In simulations, we showed that our algorithm gives better resolution compared to previously proposed blind algorithms [51, 55]. Experimentally, we demonstrated this improvement on both random speckle pattern illumination and multi-spot scanned illumination. For 3D algorithm, we quantify the resolution achieved experimentally by measuring the FWHM of the reconstructed 200 nm fluorescent beads. With the multispot illuminated SIM dataset processed by 3D algorithm, we achieved 126 nm resolution laterally and 440 nm resolution axially. In the end, we apply this algorithm on a more complicated biological sample and demonstrate 3D resolution improvement.

Chapter 4

Computational speckle structured illumination phase and fluorescence microscopy

In Chapter 2 and 3, we have explored the coherent super-resolution and incoherent super-resolution imaging techniques with the help of computational imaging approaches. Optimization relaxes the optics information retrieval process to achieve super-resolution phase retrieval without interferometric setup and super-resolution fluorescent imaging with random illumination coding. With this simplification on optical system design from the power of computational imaging, FPM demonstrates multi-fold ($4 \sim 6\times$) resolution improvement for coherent phase imaging simply by replacing the light source with an LED array. Although this resolution gain happens in an intrinsically low-NA imaging system, which does not really break the diffraction limit for high-NA system, it efficiently captures high-resolution information over large field-of-view allowed by low-NA objective, resulting in a high-content coherent imaging technique.

Here we investigate whether it is possible to make similar multi-fold resolution gains for fluorescence imaging. A key to a multi-fold resolution enhancement for fluorescence imaging is to generate an illumination pattern that has larger NA than the detection objective over a large FOV. Illumination NA determines how much super-resolution information is encoded in the system. The higher the illumination NA that is achieved, the higher resolution gain can be obtained. Larger illumination NA over a large FOV is normally difficult for traditional optical design to realize. However, being able to use random speckle illumination for super-resolution fluorescent imaging from the previous chapter has provided a simple solution to this problem, leading toward high-content fluorescence imaging. At the same time, this solution also provides a hint toward high-content coherent imaging, which pairs with fluorescence channel to form high-content multimodal imaging.

In this chapter, we focus on speckle structured illumination microscopy (SIM) as a robust and cost-effective solution for high-content fluorescence microscopy with simultaneous high-content quantitative phase (QP). This multi-modal compatibility is essential for studies

requiring cross-correlative biological analysis. Our method encodes super-resolution fluorescence and phase information by using laterally-translated Scotch tape to generate high-resolution speckle illumination patterns across a large FOV. Custom optimization algorithms then jointly reconstruct the sample’s super-resolution fluorescence (incoherent) and QP (coherent) distributions, while digitally correcting for system imperfections such as unknown speckle illumination patterns, system aberrations and pattern translations. Beyond previous linear SIM works, we achieve resolution gains of $4\times$ the objective’s diffraction-limited native resolution, resulting in 700 nm fluorescence and $1.2\ \mu\text{m}$ QP resolution, across a FOV of $2 \times 2.7\ \text{mm}^2$, giving a space-bandwidth product (SBP) of 60 megapixels.

4.1 Introduction

4.1.1 High-content microscopy

The space-bandwidth product (SBP) metric characterizes information content transmitted through an optical system; it can be thought of as the number of resolvable points in an image (*i.e.* the system’s field-of-view (FOV) divided by the size of its point spread function (PSF) [38, 132]). Typical microscopes collect images with SBPs of < 20 megapixels, a practical limit set by the systems’ optical design and camera pixel count. For large-scale biological studies in systems biology and drug discovery, fast high-SBP imaging is desired [31–37, 133]. The traditional solution for increasing SBP is to use an automated translation stage to scan the sample laterally, then stitch together high-content images. However, such capabilities are costly, have long acquisition times and require careful auto-focusing, due to small depth-of-field (DOF) and axial drift of the sample over large scan ranges [134].

Instead of using high-resolution optics and mechanically scanning the FOV, new approaches for high-content imaging use a low-NA objective (with a large FOV) and build up higher resolution by computationally combining a sequence of low-resolution measurements [17–30]. Such approaches typically illuminate the sample with customized patterns that encode high-resolution sample information into low-resolution features, which can then be measured. These methods reconstruct features smaller than the diffraction limit of the objective, using concepts from synthetic aperture [2, 11, 41] and super-resolution (SR) [3, 5–8, 58]. Though the original intent was to maximize resolution, it is important to note that by increasing resolution, SR techniques also increase SBP, and therefore have application in high-content microscopy. Eliminating the requirement for long-distance mechanical scanning means that acquisition is faster and less expensive, while focus requirements are also relaxed by the larger DOF of low-NA objectives.

Existing high-content methods generally use either an incoherent imaging model to reconstruct fluorescence [23–30], or a coherent model to reconstruct absorption and quantitative phase (QP) [17–22]. Both have achieved gigapixel-scale SBP (milli-/centi-meter scale FOV with sub-micron resolution). However, none have demonstrated cross-compatibility with both coherent (phase) and incoherent (fluorescence) imaging.

4.1.2 Contribution of this chapter

Here, we demonstrate multi-modal high-content imaging via a computational imaging framework that allows super-resolution fluorescence and QP. Our method is based on structured illumination microscopy (SIM), which is compatible with both incoherent [2, 5, 6, 60] and coherent [14, 135–139] sources of contrast [13, 16, 140, 141].

Though most SIM implementations have focused on super-resolution, some previous works have recognized its suitability for high-content imaging [23–29]. However, these predominantly relied on fluorescence imaging with calibrated illumination patterns, which are difficult to realize in practice because lens-based illumination has finite SBP. Here, we use random speckle illumination, generated by scattering through Scotch tape, in order to achieve both high-NA *and* large FOV illumination. Our method is related to blind SIM [51]; however, instead of using many random speckle patterns (which restricts resolution gain to $\sim 1.8\times$), we translate the speckle laterally, enabling resolution gains beyond that of previous methods [51, 52, 56, 57, 113, 114, 130]. Previous works also use high-cost spatial-light-modulators (SLM) [142] or galvanometer/MEMs mirrors [139, 143] for precise illumination, as well as expensive objective lenses for aberration correction. We eliminate both of these requirements by performing computational self-calibration, solving for the translation trajectory and the field-dependent aberrations of the system.

Our proposed framework enables three key advantages over existing methods:

- resolution gains of $4\times$ the native resolution of the objective (linear SIM is usually restricted to $2\times$) [51–54, 56, 57, 113, 114, 130],
- synergistic use of both the fluorescent (incoherent) and quantitative-phase (coherent) signal from the sample to enable multi-modal imaging,
- algorithmic self-calibration to significantly relax hardware requirements, enabling low-cost and robust imaging.

In our experimental setup, the Scotch tape is placed just before the sample and mounted on a translation stage (Fig. 4.1). This generates disordered speckles at the sample that are much smaller than the PSF of the imaging optics, encoding SR information. Nonlinear optimization methods are then used to jointly reconstruct multiple calibration quantities: the unknown speckle illumination pattern, the translation trajectory of the pattern, and the field-dependent system aberrations (on a patch-by-patch basis). These are subsequently used to decode the SR information of both fluorescence and phase. Compared to traditional SIM systems that use high-NA objective lenses, our system utilizes a low-NA low-cost lens to ensure large FOV. The Scotch tape generated speckle illumination is not resolution-bound by any imaging lens; this is what allows us to achieve $4\times$ resolution gains. The result is high-content imaging at sub-micron resolutions across millimeter scale regions. Various previous works have achieved cost-effectiveness, high-content (large SBP), or multiple modalities, but we believe this to be the first to simultaneously encompass all three.

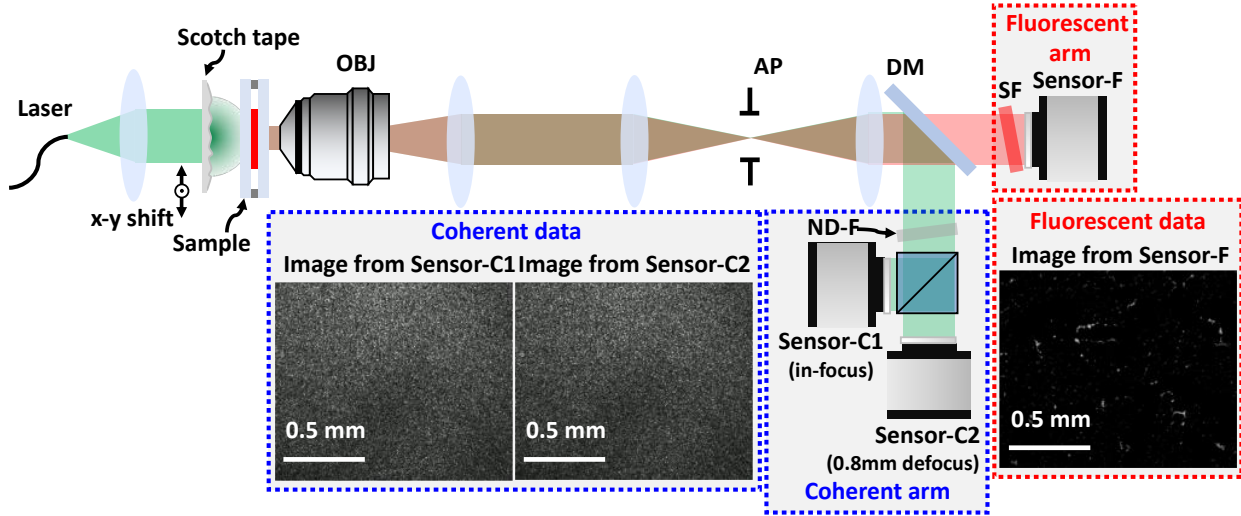


Figure 4.1: Structured illumination microscopy (SIM) with laterally-translated Scotch tape as the patterning element, achieving $4\times$ resolution gain. Our imaging system has both an incoherent arm, where Sensor-F captures raw fluorescence images (at the emission wavelength, $\lambda_{\text{em}} = 605 \text{ nm}$) for fluorescence super-resolution, and a coherent arm, where Sensor-C1 and Sensor-C2 capture images with different defocus (at the laser illumination wavelength, $\lambda_{\text{ex}} = 532 \text{ nm}$) for both super-resolution phase reconstruction and speckle trajectory calibration. OBJ: objective, AP: adjustable iris-aperture, DM: dichroic mirror, SF: spectral filter, ND-F: neutral-density filter.

4.2 Theory

SIM generally achieves super-resolution by illuminating the sample with a high spatial-frequency pattern that mixes with the sample’s information content to form low-resolution “beat” patterns (*i.e.* moire fringes). Measurements of these “beat” patterns allow elucidation of sample features beyond the diffraction-limited resolution of the imaging system. Maximum achievable resolution in SIM is set by the sum of the numerical apertures (NAs) of the illumination pattern, NA_{illum} , and the imaging system, NA_{sys} . Thus, SIM enables a resolution gain factor (over the system’s native resolution) of $(\text{NA}_{\text{illum}} + \text{NA}_{\text{sys}})/\text{NA}_{\text{sys}}$ [6]. The minimum resolvable feature size is inversely related to this bound, $d \propto 1/(\text{NA}_{\text{illum}} + \text{NA}_{\text{sys}})$.

Linear SIM typically maximizes resolution by using either: 1) a high-NA objective in epi-illumination configuration, or 2) two identical high-NA objectives in transmission geometry [6, 13]. Both result in a maximum of $2\times$ resolution gain because $\text{NA}_{\text{illum}} = \text{NA}_{\text{sys}}$, which corresponds to an SBP increase by a factor of $4\times$. Given the relatively low native SBP of high-NA imaging lenses, such increases are not sufficient to qualify as high-content imaging. Though nonlinear SIM techniques can enable higher resolution gains [58], they require either fluorophore photo-switching or saturation capabilities, which can associate with photobleaching and low SNR, and are not compatible with coherent QP techniques.

In this work, we aim for $> 2\times$ resolution gain; hence, we need the illumination NA to be larger than the detection NA, without using a high-resolution illumination lens (that would restrict the illumination FOV). To achieve this, we use a wide-area high-angle scattering element - layered Scotch tape - on the illumination side of the sample (Fig. 4.1). Multiple scattering within the tape creates a speckle pattern with finer features than the PSF of the imaging system, *i.e.* $NA_{\text{illum}} > NA_{\text{sys}}$. This means that spatial frequencies beyond $2\times$ the objective's cutoff are mixed into the measurements, which gives a chance to achieve resolution gains greater than two.

Algorithmically, to decode the high-resolution information, we need to jointly estimate the speckle pattern and the sample. This framework shares similar characteristics with the work on blind SIM first introduced by [51], where completely random speckle patterns were sequentially illuminated onto the sample. Unfortunately, the reconstruction formulation proposed in that work is especially ill-posed due to randomness between the illumination patterns, *i.e.*, if N_{img} raw images are taken, there would be $N_{\text{img}} + 1$ unknown variables to solve for (N_{img} illumination patterns and 1 sample distribution). To better condition this problem, priors based on speckle statistics [51, 56, 113, 114, 130] and sample sparsity [52, 57] can be introduced, pushing blind SIM to $2\times$ resolution gain. However, to implement high-content microscopy using SIM, we desire a resolution gain of $> 2\times$. Even with priors, we found that this degree of resolution gain was not experimentally achievable with uncorrelated and random speckle illuminations, due to the reconstruction formulation being so ill-posed.

To address this issue, we improve the posedness of the problem by illuminating with a constant speckle pattern that is translated between measurements, as opposed to randomly changing speckle patterns. Because each individual illumination pattern at the sample is a laterally shifted version of every other illumination pattern, the posedness of the reconstruction framework dramatically increases. Previous works [30, 53, 54] have also demonstrated this concept to effectively achieve beyond $2\times$ resolution gain.

The following sections outline the algorithm that we use to reconstruct large SBP fluorescence and QP images from low-resolution acquisitions of a sample illuminated by a laterally-translating speckle pattern. Our strategy relies on an inverse problem, where both the sample and illumination speckle (which is unknown) are iteratively estimated [30, 53, 54].

4.2.1 Super-resolution fluorescence imaging

Fluorescence imaging requires an incoherent imaging model. The intensity at the sensor is a low-resolution image of the sample's fluorescent distribution, obeying the system's incoherent resolution limit, $d_{\text{sys}} = \lambda_{\text{em}}/2NA_{\text{sys}}$, where λ_{em} is the emission wavelength. The speckle pattern generated through the Scotch tape excites the fluorescent sample with features of minimum size $d_{\text{illum}} = \lambda_{\text{ex}}/2NA_{\text{illum}}$, where λ_{ex} is the excitation wavelength and NA_{illum} is set by the scattering angles exiting the Scotch tape. Approximating the excitation and emission wavelengths as similar ($\lambda = \lambda_{\text{ex}} \approx \lambda_{\text{em}}$), the resolution limit of the SIM reconstruction is $d_{\text{SIM}} \approx \lambda/2(NA_{\text{sys}} + NA_{\text{illum}})$, with a resolution gain factor of $d_{\text{sys}}/d_{\text{SIM}}$. This factor is

mathematically unbounded; however, it will be practically limited by the illumination NA and SNR.

4.2.1.1 Incoherent forward model for fluorescence imaging

Plane-wave illumination of the Scotch tape, positioned at the ℓ -th scan-point, \mathbf{r}_ℓ , creates a speckle illumination pattern, $p_f(\mathbf{r} - \mathbf{r}_\ell)$, at the plane of the fluorescent sample, $o_f(\mathbf{r})$, where subscript f identifies variables in the fluorescence channel. The fluorescent signal is imaged through the system to give an intensity image at the camera plane:

$$I_{f,\ell}(\mathbf{r}) = [o_f(\mathbf{r}) \cdot \mathcal{C}\{p_f(\mathbf{r} - \mathbf{r}_\ell)\}] \otimes h_f(\mathbf{r}), \quad \ell = 1, \dots, N_{\text{img}}, \quad (4.1)$$

where \mathbf{r} is the 2D spatial coordinates (x, y) , $h_f(\mathbf{r})$ is the system PSF, and N_{img} is the total number of images captured. The subscript ℓ describes the acquisition index.

In this formulation, $o_f(\mathbf{r})$, $h_f(\mathbf{r})$, and $I_{f,\ell}(\mathbf{r})$ are 2D $M \times M$ -pixel distributions. To accurately model different regions of the pattern translating into the object's $M \times M$ FOV with incrementing \mathbf{r}_ℓ , we initialize $p_f(\mathbf{r})$ as a $N \times N$ pixel 2D distribution, with $N > M$, and introduce a cropping operator \mathcal{C} to select the $M \times M$ region of the scanning pattern that illuminates the sample.

4.2.1.2 Inverse problem for fluorescence imaging

We next formulate an optimization problem to extract SR estimates of the sample, $o_f(\mathbf{r})$, and illumination distributions, $p_f(\mathbf{r})$, from the raw fluorescence measurements, $I_{f,\ell}(\mathbf{r})$, as well as refine the estimate of the system's PSF [30] (aberrations) and speckle translation trajectory, \mathbf{r}_ℓ . We start with a crude initialization from raw observations of the speckle made using the coherent imaging arm (more details in Sec. 4.2.3). Defining $f_f(o_f, p_f, h_f, \mathbf{r}_1, \dots, \mathbf{r}_{N_{\text{img}}})$ as a joint-variable cost function that measures the difference between the raw intensity acquisitions and the expected intensities from estimated variables via the forward model, we have:

$$\min_{o_f, p_f, h_f, \mathbf{r}_1, \dots, \mathbf{r}_{N_{\text{img}}}} f_f(o_f, p_f, h_f, \mathbf{r}_1, \dots, \mathbf{r}_{N_{\text{img}}}) = \sum_{\ell=1}^{N_{\text{img}}} f_{f,\ell}(o_f, p_f, h_f, \mathbf{r}_\ell), \quad (4.2)$$

$$\text{where} \quad f_{f,\ell}(o_f, p_f, h_f, \mathbf{r}_\ell) = \sum_{\mathbf{r}} |I_{f,\ell}(\mathbf{r}) - [o_f(\mathbf{r}) \cdot \mathcal{C}\{p_f(\mathbf{r} - \mathbf{r}_\ell)\}] \otimes h_f(\mathbf{r})|^2.$$

To solve (4.2), a sequential gradient descent [144, 145] algorithm is used, where the gradient is updated once for each measurement. The sample, speckle pattern, system's PSF and scanning positions are updated by sequentially running through N_{img} measurements within one iteration. After the sequential update, an extra Nesterov's accelerated update [122] is included for both the sample and pattern estimate, to speed up convergence. Sec. 4.2.4 contains a detailed derivation of the gradient with respect to the sample, structured pattern,

system's PSF and the scanning position based on the linear algebra vectorial notation. The algorithm is described in Sec. 4.2.5.

4.2.2 Super-resolution quantitative-phase imaging

In this section, we present our coherent model for SR quantitative-phase (QP) imaging. A key difference between the QP and fluorescence imaging processes is that the detected intensity at the image plane for coherent imaging is nonlinearly related to the sample's QP [38, 136]. Thus, solving for a sample's QP from a single intensity measurement is a nonlinear and ill-posed problem. To circumvent this, we use intensity measurements from two planes, one in-focus and one out-of-focus, to introduce a complex-valued operator that couples QP variations into measurable intensity fluctuations, making the reconstruction well-posed [86, 146]. The defocused measurements are denoted by a new subscript variable z . Figure 4.1 shows our implementation, where two defocused sensors are positioned at z_0 and z_1 in the coherent imaging arm.

Generally, the resolution for coherent imaging is roughly half that of its incoherent counterpart [38]. For our QP reconstruction, the resolution limit is $d_{\text{SIM}} = \lambda_{\text{ex}} / (\text{NA}_{\text{sys}} + \text{NA}_{\text{illum}})$, where the coherent resolution of the native system and the speckle are $d_{\text{sys}} = \lambda_{\text{ex}} / \text{NA}_{\text{sys}}$ and $d_{\text{illum}} = \lambda_{\text{ex}} / \text{NA}_{\text{illum}}$, respectively.

4.2.2.1 Coherent forward model for phase imaging

Assuming an object with 2D complex transmittance function $o_c(\mathbf{r})$ is illuminated by a speckle field, $p_c(\mathbf{r})$, where subscript c refers to the coherent imaging channel, positioned at the ℓ -th scanning position \mathbf{r}_ℓ , we can represent the intensity image formed via coherent diffraction as:

$$I_{c,\ell z}(\mathbf{r}) = |[o_c(\mathbf{r}) \cdot \mathcal{C}\{p_c(\mathbf{r} - \mathbf{r}_\ell)\}] \otimes h_{c,z}(\mathbf{r})|^2 = |g_{c,\ell z}(\mathbf{r})|^2, \ell = 1, \dots, N_{\text{img}}, z = z_0, z_1, (4.3)$$

where $g_{c,\ell z}(\mathbf{r})$ and $h_{c,z}(\mathbf{r})$ are the complex electric-fields at the imaging plane and the system's coherent PSF at defocus distance z , respectively. The comma in the subscript separates the channel index, c or f , from the scanning-position and acquisition-number indices, ℓ and z . N_{img} here indicates the total number of translations of the Scotch tape. The defocused PSF can be further broken down into $h_{c,z}(\mathbf{r}) = h_c(\mathbf{r}) \otimes h_z(\mathbf{r})$, where $h_c(\mathbf{r})$ is the in-focus coherent PSF and $h_z(\mathbf{r})$ is the defocus kernel. Similar to Section 4.2.1.1, $I_{c,\ell z}(\mathbf{r})$, $o_c(\mathbf{r})$, and $h_{c,z}(\mathbf{r})$ are 2D distributions with dimensions of $M \times M$ pixels, while $p_c(\mathbf{r})$ is of size $N \times N$ pixels ($N > M$). \mathcal{C} is a cropping operator that selects the sub-region of the pattern that interacts with the sample. The sample's QP distribution is simply the phase of the object's complex transmittance, $\angle o_c(\mathbf{r})$.

4.2.2.2 Inverse problem for phase imaging

We now take the raw coherent intensity measurements, $I_{c,\ell z}(\mathbf{r})$, and the registered trajectory, $\mathbf{r}_{\ell z}$, from both of the defocused coherent sensors (more details in Sec. 4.2.3) as input to jointly

estimate the sample's SR complex-transmittance function, $o_c(\mathbf{r})$, and illumination complex-field, $p_c(\mathbf{r})$, as well as the aberrations inherent in the system's PSF, $h_c(\mathbf{r})$. The optimization also further refines the scanning trajectory, $\mathbf{r}_{\ell z}$. Based on the forward model, we formulate the joint inverse problem:

$$\begin{aligned} & \underset{\substack{o_c, p_c, h_c, \mathbf{r}_{1z_0}, \mathbf{r}_{1z_1}, \\ \dots, \mathbf{r}_{N_{\text{img}}z_0}, \mathbf{r}_{N_{\text{img}}z_1}}}{\text{minimize}} \quad f_c(o_c, p_c, h_c, \mathbf{r}_{1z_0}, \mathbf{r}_{1z_1}, \dots, \mathbf{r}_{N_{\text{img}}z_0}, \mathbf{r}_{N_{\text{img}}z_1}) = \sum_{\ell, z} f_{c, \ell z}(o_c, p_c, h_c, \mathbf{r}_{\ell z}), \\ \text{where} \quad & f_{c, \ell z}(o_c, p_c, h_c, \mathbf{r}_{\ell z}) = \sum_{\mathbf{r}} \left| \sqrt{I_{c, \ell z}(\mathbf{r})} - |[o_c(\mathbf{r}) \cdot \mathcal{C}\{p_c(\mathbf{r} - \mathbf{r}_{\ell z})\}] \otimes h_{c, z}(\mathbf{r})| \right|^2. \end{aligned} \quad (4.4)$$

Here, we adopt an amplitude-based cost function, f_c , which robustly minimizes the distance between the estimated and measured amplitudes in the presence of noise [84, 86, 144]. We optimize the pattern trajectories, \mathbf{r}_{ℓ, z_0} and \mathbf{r}_{ℓ, z_1} , separately for each coherent sensor, in order to account for any residual misalignment or timing-mismatch (see Sec. 4.2.3). As in the fluorescence case, sequential gradient descent [144, 145] was used to solve this inverse problem.

4.2.3 Registration of coherent images

Knowledge of the Scotch tape scanning position, \mathbf{r}_ℓ , reduces the complexity of the joint sample and pattern estimation problem and is necessary to achieve SR reconstructions with greater than $2\times$ resolution gain. Because our fluorescent sample is mostly transparent, the main scattering component in the acquired raw data originates from the Scotch tape. Thus, using a sub-pixel registration algorithm [147] between successive coherent-camera acquisitions, which are dominated by the scattered speckle signal, is sufficient to initialize the scanning trajectory of the Scotch tape,

$$\mathbf{r}_{\ell z} = \mathcal{R}[I_{c, 1z}(\mathbf{r}), I_{c, \ell z}(\mathbf{r})], \quad (4.5)$$

where \mathcal{R} is the registration operator. These initial estimates of $\mathbf{r}_{\ell z}$ are then updated, alongside $o_f(\mathbf{r})$, $o_c(\mathbf{r})$, $p_f(\mathbf{r})$, and $p_c(\mathbf{r})$ using the inverse models described in Sec. 4.2.1.2 and 4.2.2.2. In the fluorescence problem described in Sec. 4.2.1.2, we only use the trajectory from the in-focus coherent sensor at $z = 0$ for initialization, so we omit the subscript z in $\mathbf{r}_{\ell z}$.

4.2.4 Solving the inverse problem

4.2.4.1 Vectorial notation

Fluorescence imaging vectorial model

In order to solve the multivariate optimization problem in Eq. (4.2) and (4.4) and derive the gradient of the cost function, it is more convenient to consider a linear algebra vectorial

notation of the forward models. The fluorescence SIM forward model in Eq. (4.1) can be alternatively expressed as

$$\mathbf{I}_{f,\ell} = \mathbf{H}_f \text{diag}(\mathbf{S}(\mathbf{r}_\ell) \mathbf{p}_f) \mathbf{o}_f, \quad (4.6)$$

where $\mathbf{I}_{f,\ell}$, \mathbf{H}_f , $\mathbf{S}(\mathbf{r}_\ell)$, \mathbf{p}_f , and \mathbf{o}_f designate the raw fluorescent intensity vector, diffraction-limit low-pass filtering operation, pattern translation/cropping operation, $N^2 \times 1$ speckle pattern vector, and $M^2 \times 1$ sample's fluorescent distribution vector, respectively. The 2D-array variables described in (4.1) are all reshaped into column vectors here. \mathbf{H}_f and $\mathbf{S}(\mathbf{r}_\ell)$ can be further broken down into their individual vectorial components:

$$\begin{aligned} \mathbf{H}_f &= \mathbf{F}_M^{-1} \text{diag}(\tilde{\mathbf{h}}_f) \mathbf{F}_M, \\ \mathbf{S}(\mathbf{r}_\ell) &= \mathbf{Q} \mathbf{F}_N^{-1} \text{diag}(\mathbf{e}(\mathbf{r}_\ell)) \mathbf{F}_N, \end{aligned} \quad (4.7)$$

where $\tilde{\mathbf{h}}_f$ is the OTF vector and $\mathbf{e}(\mathbf{r}_\ell)$ is the vectorization of the $\exp(-j2\pi \mathbf{u} \cdot \mathbf{r}_\ell)$ function, where \mathbf{u} is spatial frequency. The notation $\text{diag}(\mathbf{a})$ turns a $n \times 1$ vector, \mathbf{a} , into an $n \times n$ diagonal matrix with diagonal entries from the vector entries. \mathbf{F}_N and \mathbf{F}_M denote the $N \times N$ -point and $M \times M$ -point 2D discrete Fourier transform matrix, respectively, and \mathbf{Q} is the $M^2 \times N^2$ cropping matrix.

With this vectorial notation, the cost function for a single fluorescence measurement is

$$f_{f,\ell}(\mathbf{o}_f, \mathbf{p}_f, \tilde{\mathbf{h}}_f, \mathbf{r}_\ell) = \mathbf{f}_{f,\ell}^T \mathbf{f}_{f,\ell} = \|\mathbf{I}_{f,\ell} - \mathbf{H}_f \text{diag}(\mathbf{S}(\mathbf{r}_\ell) \mathbf{p}_f) \mathbf{o}_f\|_2^2, \quad (4.8)$$

where $\mathbf{f}_{f,\ell} = \mathbf{I}_{f,\ell} - \mathbf{H}_f \text{diag}(\mathbf{S}(\mathbf{r}_\ell) \mathbf{p}_f) \mathbf{o}_f$ is the cost vector and T denotes the transpose operation.

Coherent imaging vectorial model

As with the fluorescence vectorial model, we can rewrite Eq. (4.3) using vectorial notation:

$$\mathbf{I}_{c,lz} = |\mathbf{g}_{c,lz}|^2, \quad (4.9)$$

where

$$\begin{aligned} \mathbf{g}_{c,lz} &= \mathbf{H}_{c,z} \text{diag}(\mathbf{S}(\mathbf{r}_{lz}) \mathbf{p}_c) \mathbf{o}_c \\ \mathbf{H}_{c,z} &= \mathbf{F}_M^{-1} \text{diag}(\tilde{\mathbf{h}}_c) \text{diag}(\tilde{\mathbf{h}}_z) \mathbf{F}_M. \end{aligned} \quad (4.10)$$

\mathbf{o}_c and \mathbf{p}_c are the $M^2 \times 1$ sample transmittance function vector and $N^2 \times 1$ structured field vector, respectively. $\tilde{\mathbf{h}}_c$ and $\tilde{\mathbf{h}}_z$ are the system pupil function and the deliberate defocus pupil function, respectively. With this vectorial notation, we can then express the cost function for a single coherent intensity measurement as

$$f_{c,lz}(\mathbf{o}_c, \mathbf{p}_c, \tilde{\mathbf{h}}_c, \mathbf{r}_{lz}) = \mathbf{f}_{c,lz}^T \mathbf{f}_{c,lz} = \left\| \sqrt{\mathbf{I}_{c,lz}} - |\mathbf{g}_{c,lz}| \right\|_2^2, \quad (4.11)$$

where $\mathbf{f}_{c,lz} = \sqrt{\mathbf{I}_{c,lz}} - |\mathbf{g}_{c,lz}|$ is the cost vector for the coherent intensity measurement.

4.2.4.2 Gradient derivation

Gradient derivation for fluorescence imaging

To optimize Eq. (4.2) for the variables \mathbf{o}_f , \mathbf{p}_f , $\tilde{\mathbf{h}}_f$ and \mathbf{r}_ℓ , we first derive the necessary gradients of the fluorescence cost function. Consider taking the gradient of $f_{f,\ell}$ with respect to \mathbf{o}_f , we can represent the $1 \times M^2$ gradient row vector as

$$\frac{\partial f_{f,\ell}}{\partial \mathbf{o}_f} = \left(\frac{\partial f_{f,\ell}}{\partial \mathbf{f}_{f,\ell}} \right) \cdot \left(\frac{\partial \mathbf{f}_{f,\ell}}{\partial \mathbf{o}_f} \right) = (2\mathbf{f}_{f,\ell}^T) \cdot (-\mathbf{H}_f \text{diag}(\mathbf{S}(\mathbf{r}_\ell)\mathbf{p}_f)). \quad (4.12)$$

Turning the row gradient vector into a $M^2 \times 1$ column vector in order to update the object vector in the right dimension, we the final gradient becomes

$$\nabla_{\mathbf{o}_f} f_{f,\ell} = \left(\frac{\partial f_{f,\ell}}{\partial \mathbf{o}_f} \right)^T = -2 \text{diag}(\mathbf{S}(\mathbf{r}_\ell)\mathbf{p}_f) \mathbf{H}_f^T \mathbf{f}_{f,\ell}. \quad (4.13)$$

To compute the gradient of \mathbf{p}_f , we first rewrite the cost vector $\mathbf{f}_{f,\ell}$ as

$$\mathbf{f}_{f,\ell} = \mathbf{I}_{f,\ell} - \mathbf{H}_f \text{diag}(\mathbf{o}) \mathbf{S}(\mathbf{r}_\ell)\mathbf{p}_f. \quad (4.14)$$

Now, we can write the gradient of the cost function with respect to the pattern vector in row and column vector form as

$$\begin{aligned} \frac{\partial f_{f,\ell}}{\partial \mathbf{p}_f} &= \left(\frac{\partial f_{f,\ell}}{\partial \mathbf{f}_{f,\ell}} \right) \cdot \left(\frac{\partial \mathbf{f}_{f,\ell}}{\partial \mathbf{p}_f} \right) = (2\mathbf{f}_{f,\ell}^T) \cdot (-\mathbf{H}_f \text{diag}(\mathbf{o}_f) \mathbf{S}(\mathbf{r}_\ell)) \\ \nabla_{\mathbf{p}_f} f_{f,\ell} &= \left(\frac{\partial f_{f,\ell}}{\partial \mathbf{p}_f} \right)^T = -2\mathbf{S}(\mathbf{r}_\ell)^T \text{diag}(\mathbf{o}_f) \mathbf{H}_f^T \mathbf{f}_{f,\ell}. \end{aligned} \quad (4.15)$$

Similar to the derivation of the pattern function gradient, it is easier to work with the rewritten form of the cost vector expressed as

$$\mathbf{f}_{f,\ell} = \mathbf{I}_{f,\ell} - \mathbf{F}_M^{-1} \text{diag}(\mathbf{F}_M \text{diag}(\mathbf{S}(\mathbf{r}_\ell)\mathbf{p}_f\mathbf{o}_f)) \tilde{\mathbf{h}}_f. \quad (4.16)$$

The gradient of the cost function with respect to the OTF vector in the row and column vector form are expressed, respectively, as

$$\begin{aligned} \frac{\partial f_{f,\ell}}{\partial \tilde{\mathbf{h}}_f} &= \left(\frac{\partial f_{f,\ell}}{\partial \mathbf{f}_{f,\ell}} \right) \cdot \left(\frac{\partial \mathbf{f}_{f,\ell}}{\partial \tilde{\mathbf{h}}_f} \right) = (2\mathbf{f}_{f,\ell}^T) \cdot (-\mathbf{F}_M^{-1} \text{diag}(\mathbf{F}_M \text{diag}(\mathbf{S}(\mathbf{r}_\ell)\mathbf{p}_f\mathbf{o}_f))) \\ \nabla_{\tilde{\mathbf{h}}_f} f_{f,\ell} &= \left(\frac{\partial f_{f,\ell}}{\partial \tilde{\mathbf{h}}_f} \right)^\dagger = -2 \text{diag}(\overline{\mathbf{F}_M \text{diag}(\mathbf{S}(\mathbf{r}_\ell)\mathbf{p}_f\mathbf{o}_f)}) \mathbf{F}_M \mathbf{f}_{f,\ell}, \end{aligned} \quad (4.17)$$

where $\bar{\mathbf{a}}$ denotes entry-wise complex conjugate operation on any general vector \mathbf{a} . One difference between this gradient and the previous one is that the variable to solve, $\tilde{\mathbf{h}}_f$, is now

a complex vector. When turning the gradient row vector of a complex vector into a column vector, we have to take a Hermitian operation, \dagger , on the row vector following the conventions in [103]. We will have more examples of complex variables in the coherent model gradient derivation.

For taking the gradient of the scanning position, we again rewrite the cost vector $\mathbf{f}_{f,\ell}$:

$$\mathbf{f}_{f,\ell} = \mathbf{I}_\ell - \mathbf{H}_f \text{diag}(\mathbf{o}_f) \mathbf{Q} \mathbf{F}_N^{-1} \text{diag}(\mathbf{F}_N \mathbf{p}_f) \mathbf{e}(\mathbf{r}_\ell). \quad (4.18)$$

We can then write the gradient of the cost function with respect to the scanning position as

$$\begin{aligned} \frac{\partial f_{f,\ell}}{\partial q_\ell} &= \left(\frac{\partial f_{f,\ell}}{\partial \mathbf{f}_{f,\ell}} \right) \cdot \left(\frac{\partial \mathbf{f}_{f,\ell}}{\partial \mathbf{e}(\mathbf{r}_\ell)} \right) \cdot \left(\frac{\partial \mathbf{e}(\mathbf{r}_\ell)}{\partial q_\ell} \right) \\ &= (2\mathbf{f}_{f,\ell}^T) \cdot (-\mathbf{H}_f \text{diag}(\mathbf{o}_f) \mathbf{Q} \mathbf{F}_N^{-1} \text{diag}(\mathbf{F}_N \mathbf{p}_f)) \cdot (\text{diag}(-j2\pi \mathbf{u}_q) \mathbf{e}(\mathbf{r}_\ell)), \end{aligned} \quad (4.19)$$

where q is either the x or y spatial coordinate component of \mathbf{r}_ℓ . \mathbf{u}_q is the $N^2 \times 1$ vectorial notation of the spatial frequency function in the q direction.

To numerically evaluate these gradients, we represent them in the functional form as:

$$\begin{aligned} \nabla_{o_f} f_{f,\ell}(o_f, p_f, h_f, \mathbf{r}_\ell) &= -2p_f(\mathbf{r} - \mathbf{r}_\ell) \cdot [h_f^*(-\mathbf{r}) \otimes (I_{f,\ell}(\mathbf{r}) - [o_f(\mathbf{r}) \cdot \mathcal{C}\{p_f(\mathbf{r} - \mathbf{r}_\ell)\}] \otimes h_f(\mathbf{r}))], \\ \nabla_{p_f} f_{f,\ell}(o_f, p_f, h_f, \mathbf{r}_\ell) &= -2\delta(\mathbf{r} + \mathbf{r}_\ell) \otimes \mathcal{P} \{o_f(\mathbf{r}) \cdot [h_f^*(-\mathbf{r}) \otimes (I_{f,\ell}(\mathbf{r}) - [o_f(\mathbf{r}) \cdot \mathcal{C}\{p_f(\mathbf{r} - \mathbf{r}_\ell)\}] \otimes h_f(\mathbf{r}))]\}, \\ \nabla_{\tilde{h}_f} f_{f,\ell}(o_f, p_f, h_f, \mathbf{r}_\ell) &= -2(\mathcal{F} \{o_f(\mathbf{r}) \cdot \mathcal{C}\{p_f(\mathbf{r} - \mathbf{r}_\ell)\}\})^* \cdot \mathcal{F} \{I_{f,\ell}(\mathbf{r}) - [o_f(\mathbf{r}) \cdot \mathcal{C}\{p_f(\mathbf{r} - \mathbf{r}_\ell)\}] \otimes h_f(\mathbf{r})\}, \\ \nabla_{q_\ell} f_{f,\ell}(o_f, p_f, h_f, \mathbf{r}_\ell) &= -2 \left\{ \sum_{\mathbf{r}} (I_{f,\ell}(\mathbf{r}) - [o_f(\mathbf{r}) \cdot \mathcal{C}\{p_f(\mathbf{r} - \mathbf{r}_\ell)\}] \otimes h_f(\mathbf{r})) \cdot \right. \\ &\quad \left. h_f(\mathbf{r}) \otimes \left[o_f(\mathbf{r}) \cdot \mathcal{C} \left\{ \frac{\partial p_f(\mathbf{r} - \mathbf{r}_\ell)}{\partial q_\ell} \right\} \right] \right\}, \end{aligned} \quad (4.20)$$

where a^* stands for complex conjugate of function, a , \mathcal{F} is the Fourier transform operator, and \mathcal{P} is a zero-padding operator that pads an $M \times M$ image to size $N \times N$ pixels. In this form, $I_{f,\ell}(\mathbf{r})$, $o_f(\mathbf{r})$, and $h_f(\mathbf{r})$ are 2D $M \times M$ images, while $p_f(\mathbf{r})$ is a $N \times N$ image. The gradients for the sample and the structured pattern are of the same size as $o_f(\mathbf{r})$ and $p_f(\mathbf{r})$, respectively. Ideally, the gradient of the the scanning position in each direction is a real number. However, due to imperfect implementation of the discrete differentiation in each direction, the gradient will have small imaginary value that will be dropped in the update of the scanning position.

Gradient derivation for coherent imaging

For the coherent imaging case, we will derive the gradients of the cost function in Eq. (4.11) with respect to the sample transmittance function \mathbf{o}_c , speckle field \mathbf{p}_c , pupil function $\tilde{\mathbf{h}}_c$, and the scanning position $\mathbf{r}_{\ell z}$. First, we take the gradient of $f_{c,\ell z}$ with respect to \mathbf{o}_c , we then

have the gradient in the row and column vector forms as

$$\begin{aligned}
 \frac{\partial f_{c,lz}}{\partial \mathbf{o}_c} &= \left(\frac{\partial f_{c,lz}}{\partial \mathbf{f}_{c,lz}} \right) \cdot \left(\frac{\partial \mathbf{f}_{c,lz}}{\partial \mathbf{g}_{c,lz}} \right) \cdot \left(\frac{\partial \mathbf{g}_{c,lz}}{\partial \mathbf{o}_c} \right) \\
 &= (2\mathbf{f}_{c,lz}^T) \cdot \left(-\frac{1}{2} \text{diag} \left(\frac{\overline{\mathbf{g}_{c,lz}}}{|\mathbf{g}_{c,lz}|} \right) \right) \cdot (\mathbf{H}_{c,z} \text{diag}(\mathbf{S}(\mathbf{r}_{lz})) \mathbf{p}_c) \\
 \nabla_{\mathbf{o}_c} f_{c,lz} &= \left(\frac{\partial f_{c,lz}}{\partial \mathbf{o}_c} \right)^\dagger = -\text{diag}(\overline{\mathbf{S}(\mathbf{r}_{lz}) \mathbf{p}_c}) \mathbf{H}_{c,z}^\dagger \text{diag} \left(\frac{\mathbf{g}_{c,lz}}{|\mathbf{g}_{c,lz}|} \right) \mathbf{f}_{c,lz}, \quad (4.21)
 \end{aligned}$$

where the $\frac{\mathbf{g}_{c,lz}}{|\mathbf{g}_{c,lz}|}$ operation denotes entry-wise division between the two vectors, $\mathbf{g}_{c,lz}$ and $|\mathbf{g}_{c,lz}|$. In addition, the detailed calculation of $\frac{\partial \mathbf{f}_{c,lz}}{\partial \mathbf{g}_{c,lz}}$ can be found in the Appendix of [144]. Next, we take the gradient with respect to the pattern field vector, \mathbf{p}_c , and write down the corresponding row and column vectors as

$$\begin{aligned}
 \frac{\partial f_{c,lz}}{\partial \mathbf{p}_c} &= \left(\frac{\partial f_{c,lz}}{\partial \mathbf{f}_{c,lz}} \right) \cdot \left(\frac{\partial \mathbf{f}_{c,lz}}{\partial \mathbf{g}_{c,lz}} \right) \cdot \left(\frac{\partial \mathbf{g}_{c,lz}}{\partial \mathbf{p}_c} \right) \\
 &= (2\mathbf{f}_{c,lz}^T) \cdot \left(-\frac{1}{2} \text{diag} \left(\frac{\overline{\mathbf{g}_{c,lz}}}{|\mathbf{g}_{c,lz}|} \right) \right) \cdot (\mathbf{H}_{c,z} \text{diag}(\mathbf{o}_c) \mathbf{S}(\mathbf{r}_{lz})) \\
 \nabla_{\mathbf{p}_c} f_{c,lz} &= \left(\frac{\partial f_{c,lz}}{\partial \mathbf{p}_c} \right)^\dagger = -\mathbf{S}(\mathbf{r}_{lz})^\dagger \text{diag}(\overline{\mathbf{o}_c}) \mathbf{H}_{c,z}^\dagger \text{diag} \left(\frac{\mathbf{g}_{c,lz}}{|\mathbf{g}_{c,lz}|} \right) \mathbf{f}_{c,lz}. \quad (4.22)
 \end{aligned}$$

In order to calculate $\frac{\partial \mathbf{g}_{c,lz}}{\partial \mathbf{p}_c}$, we need to reorder the dot multiplication of \mathbf{o}_c and $\mathbf{S}(\mathbf{r}_{lz}) \mathbf{p}_c$ as we did in deriving the gradient of the pattern for fluorescence imaging.

In order to do aberration correction, we will need to estimate the system pupil function, $\tilde{\mathbf{h}}_c$. The gradient with respect to the pupil function can be derived as,

$$\begin{aligned}
 \frac{\partial f_{c,lz}}{\partial \tilde{\mathbf{h}}_c} &= \left(\frac{\partial f_{c,lz}}{\partial \mathbf{f}_{c,lz}} \right) \cdot \left(\frac{\partial \mathbf{f}_{c,lz}}{\partial \mathbf{g}_{c,lz}} \right) \cdot \left(\frac{\partial \mathbf{g}_{c,lz}}{\partial \tilde{\mathbf{h}}_c} \right) \\
 &= (2\mathbf{f}_{c,lz}^T) \cdot \left(-\frac{1}{2} \text{diag} \left(\frac{\overline{\mathbf{g}_{c,lz}}}{|\mathbf{g}_{c,lz}|} \right) \right) \cdot \left(\mathbf{F}_M^{-1} \text{diag} [\mathbf{F}_M \text{diag}(\mathbf{S}(\mathbf{r}_{lz}) \mathbf{p}_c) \mathbf{o}_c] \text{diag}(\tilde{\mathbf{h}}_z) \right) \\
 \nabla_{\tilde{\mathbf{h}}_c} f_{c,lz} &= \left(\frac{\partial f_{c,lz}}{\partial \tilde{\mathbf{h}}_c} \right)^\dagger = -\text{diag}(\overline{\tilde{\mathbf{h}}_z}) \text{diag} \left[\overline{\mathbf{F}_M \text{diag}(\mathbf{S}(\mathbf{r}_{lz}) \mathbf{p}_c) \mathbf{o}_c} \right] \mathbf{F}_M \text{diag} \left(\frac{\mathbf{g}_{c,lz}}{|\mathbf{g}_{c,lz}|} \right) \mathbf{f}_{c,lz}. \quad (4.23)
 \end{aligned}$$

In the end, the gradient of the scanning position for refinement can be derived as

$$\begin{aligned}
 \frac{\partial f_{c,\ell z}}{\partial q_{\ell z}} &= \left(\frac{\partial f_{c,\ell z}}{\partial \mathbf{f}_{c,\ell z}} \right) \cdot \left[\left(\frac{\partial \mathbf{f}_{c,\ell z}}{\partial \mathbf{g}_{c,\ell z}} \right) \cdot \left(\frac{\partial \mathbf{g}_{c,\ell z}}{\partial \mathbf{e}(\mathbf{r}_{\ell z})} \right) \cdot \left(\frac{\mathbf{e}(\mathbf{r}_{\ell z})}{\partial q_{\ell}} \right) + \left(\frac{\partial \mathbf{f}_{c,\ell z}}{\partial \overline{\mathbf{g}_{c,\ell z}}} \right) \cdot \left(\frac{\overline{\mathbf{e}(\mathbf{r}_{\ell z})}}{\partial q_{\ell}} \right) \right] \\
 &= 2 \left(\frac{\partial f_{c,\ell z}}{\partial \mathbf{f}_{c,\ell z}} \right) \cdot \text{Re} \left\{ \left(\frac{\partial \mathbf{f}_{c,\ell z}}{\partial \mathbf{g}_{c,\ell z}} \right) \cdot \left(\frac{\partial \mathbf{g}_{c,\ell z}}{\partial \mathbf{e}(\mathbf{r}_{\ell z})} \right) \cdot \left(\frac{\mathbf{e}(\mathbf{r}_{\ell z})}{\partial q_{\ell}} \right) \right\} \\
 &= 2 \left(2\mathbf{f}_{c,\ell z}^T \right) \cdot \text{Re} \left\{ \left(-\frac{1}{2} \text{diag} \left(\frac{\overline{\mathbf{g}_{c,\ell z}}}{|\mathbf{g}_{c,\ell z}|} \right) \right) \cdot \left(\mathbf{H}_{c,z} \text{diag}(\mathbf{o}_c) \mathbf{Q} \mathbf{F}_N^{-1} \text{diag}(\mathbf{F}_N \mathbf{p}_c) \right) \cdot \left(\text{diag}(-j2\pi \mathbf{u}_q) \mathbf{e}(\mathbf{r}_{\ell z}) \right) \right\} \\
 &= -2 \text{Re} \left\{ \mathbf{f}_{c,\ell z}^T \text{diag} \left(\frac{\overline{\mathbf{g}_{c,\ell z}}}{|\mathbf{g}_{c,\ell z}|} \right) \mathbf{H}_{c,z} \text{diag}(\mathbf{o}_c) \mathbf{Q} \mathbf{F}_N^{-1} \text{diag}(\mathbf{F}_N \mathbf{p}_c) \text{diag}(-j2\pi \mathbf{u}_q) \mathbf{e}(\mathbf{r}_{\ell z}) \right\}, \tag{4.24}
 \end{aligned}$$

where q is either the x or y spatial coordinate component of $\mathbf{r}_{\ell z}$.

In order to numerically evaluate these gradients, we represent them, as we did for the gradients of the fluorescence model, into functional forms:

$$\begin{aligned}
 \nabla_{o_c} f_{c,\ell z}(o_c, p_c, h_c, \mathbf{r}_{\ell z}) &= -p_c^*(\mathbf{r} - \mathbf{r}_{\ell z}) \cdot \left[h_{c,z}^*(-\mathbf{r}) \otimes \left(\left(\frac{\sqrt{I_{c,\ell z}(\mathbf{r})}}{|g_{c,\ell z}(\mathbf{r})|} - 1 \right) \cdot g_{c,\ell z}(\mathbf{r}) \right) \right] \\
 \nabla_{p_c} f_{c,\ell z}(o_c, p_c, h_c, \mathbf{r}_{\ell z}) &= -\delta(\mathbf{r} + \mathbf{r}_{\ell z}) \otimes \mathcal{P} \left\{ o_c^*(\mathbf{r}) \cdot \left[h_{c,z}^*(-\mathbf{r}) \otimes \left(\left(\frac{\sqrt{I_{c,\ell z}(\mathbf{r})}}{|g_{c,\ell z}(\mathbf{r})|} - 1 \right) \cdot g_{c,\ell z}(\mathbf{r}) \right) \right] \right\} \\
 \nabla_{\tilde{h}_c} f_{c,\ell z}(o_c, p_c, h_c, \mathbf{r}_{\ell z}) &= -\tilde{h}_z^*(\mathbf{u}) \cdot \mathcal{F} \{ p_c(\mathbf{r} - \mathbf{r}_{\ell z}) \cdot o_c(\mathbf{r}) \}^* \mathcal{F} \left\{ \left(\frac{\sqrt{I_{c,\ell z}(\mathbf{r})}}{|g_{c,\ell z}(\mathbf{r})|} - 1 \right) \cdot g_{c,\ell z}(\mathbf{r}) \right\} \\
 \nabla_{q_{\ell z}} f_{c,\ell z}(o_c, p_c, h_c, \mathbf{r}_{\ell z}) &= -2 \text{Re} \left\{ \sum_{\mathbf{r}} \left[\left(\frac{\sqrt{I_{c,\ell z}(\mathbf{r})}}{|g_{c,\ell z}(\mathbf{r})|} - 1 \right) \cdot g_{c,\ell z}^*(\mathbf{r}) \right] \cdot \left[h_{c,z}(\mathbf{r}) \otimes \left(o_c(\mathbf{r}) \cdot \mathcal{C} \left\{ \frac{\partial p_c(\mathbf{r} - \mathbf{r}_{\ell z})}{\partial q_{\ell z}} \right\} \right) \right] \right\}. \tag{4.25}
 \end{aligned}$$

4.2.5 Reconstruction algorithm

With the derivation of the gradients in Sec. 4.2.4, we summarize here the reconstruction algorithm for fluorescence imaging and coherent imaging.

4.2.5.1 Algorithm for fluorescent imaging

First, we initialize the sample, $o_f(\mathbf{r})$, with the mean image of all the structure illuminated images, $I_{f,\ell}(\mathbf{r})$, which is approximately a widefield diffraction-limited image. As for the structured pattern, $p_f(\mathbf{r})$, we initialize it with a all-one image. The initial OTF, $\tilde{h}_f(\mathbf{u})$, is set as a non-aberrated incoherent OTF. Initial scanning positions are from the registration of the in-focus coherent speckle images, $I_{c,\ell z}(\mathbf{r})$ ($z = 0$).

In the algorithm, K_f is the total number of iterations ($K_f = 100$ is generally enough for convergence). At every iteration, we sequentially update the sample, structured pattern, system's OTF and the scanning position using each single frame from $\ell = 1$ to $\ell = N_{\text{img}}$. A Nesterov acceleration step is applied on the sample and the structured pattern at the end of each iteration. The detailed algorithm is summarized in Algorithm 5.

Algorithm 2 Fluorescence imaging reconstruction

Require: $I_{f,\ell}(\mathbf{r})$, \mathbf{r}_ℓ , $\ell = 1, \dots, N_{\text{img}}$

- 1: initialize $o_f^{(1,0)}(\mathbf{r}) = \sum_{\ell} I_{f,\ell}(\mathbf{r})/N_{\text{img}}$
 - 2: initialize $p_f^{(1,0)}(\mathbf{r})$ with all one values
 - 3: initialize $\tilde{h}_f(\mathbf{u})$ with the non-aberrated incoherent OTF
 - 4: initialize $\mathbf{r}_\ell^{(1)}$ with the scanning position from the registration step
 - 5: **for** $k = 1 : K_f$ **do**
 - 6: Sequential gradient descent
 - 7: **for** $\ell = 1 : N_{\text{img}}$ **do**
 - 8: $o_f^{(k,\ell)}(\mathbf{r}) = o_f^{(k,\ell-1)}(\mathbf{r}) - \nabla_{o_f} f_{f,\ell}(o_f^{(k,\ell-1)}, p_f^{(k,\ell-1)}, \mathbf{r}_\ell^{(k)}) / \max(p_f^{(k,\ell-1)}(\mathbf{r}))^2$
 - 9: $p_f^{(k,\ell)}(\mathbf{r}) = p_f^{(k,\ell-1)}(\mathbf{r}) - \nabla_{p_f} f_{f,\ell}(o_f^{(k,\ell-1)}, p_f^{(k,\ell-1)}, \mathbf{r}_\ell^{(k)}) / \max(o_f^{(k,\ell-1)}(\mathbf{r}))^2$
 - 10: $\xi(\mathbf{u}) = \mathcal{F}\{o_f^{(k,\ell-1)}(\mathbf{r}) \cdot \mathcal{C}\{p_f^{(k,\ell-1)}(\mathbf{r} - \mathbf{r}_\ell)\}\}$
 - 11: $\tilde{h}_f^{(k,\ell)}(\mathbf{u}) = \tilde{h}_f^{(k,\ell-1)}(\mathbf{u}) - \nabla_{\tilde{h}_f} f_{f,\ell}(o_f^{(k,\ell-1)}, p_f^{(k,\ell-1)}, h_f^{(k,\ell-1)}, \mathbf{r}_\ell^{(k)}) \cdot |\xi(\mathbf{u})| / 12[\max(|\xi(\mathbf{u})| \cdot (|\xi(\mathbf{u})|^2 + \delta))]$, where δ is chosen to be small
 - 12:
 - 13: Scanning position refinement
 - 14: $x_\ell^{(k+1)} = x_\ell^{(k)} - \alpha \nabla_{x_\ell} f_{f,\ell}(o_f^{(k,\ell-1)}, p_f^{(k,\ell-1)}, \mathbf{r}_\ell^{(k)})$
 - 15: $y_\ell^{(k+1)} = y_\ell^{(k)} - \alpha \nabla_{y_\ell} f_{f,\ell}(o_f^{(k,\ell-1)}, p_f^{(k,\ell-1)}, \mathbf{r}_\ell^{(k)})$
 - 16: **end for**
 - 17: Nesterov's acceleration
 - 18: **if** $k = 1$ **then**
 - 19: $t_1 = 1$
 - 20: $o_f^{(k+1,0)}(\mathbf{r}) = o_f^{(k,N_{\text{img}})}(\mathbf{r})$
 - 21: $p_f^{(k+1,0)}(\mathbf{r}) = p_f^{(k,N_{\text{img}})}(\mathbf{r})$
 - 22: **else**
 - 23: $t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2}$
 - 24: $o_f^{(k+1,0)}(\mathbf{r}) = o_f^{(k,N_{\text{img}})}(\mathbf{r}) + \frac{t_k - 1}{t_{k+1}} \left[o_f^{(k,N_{\text{img}})}(\mathbf{r}) - o_f^{(k-1,N_{\text{img}})}(\mathbf{r}) \right]$
 - 25: $p_f^{(k+1,0)}(\mathbf{r}) = p_f^{(k,N_{\text{img}})}(\mathbf{r}) + \frac{t_k - 1}{t_{k+1}} \left[p_f^{(k,N_{\text{img}})}(\mathbf{r}) - p_f^{(k-1,N_{\text{img}})}(\mathbf{r}) \right]$
 - 26: **end if**
 - 27: **end for**
-

4.2.5.2 Algorithm for coherent imaging

For coherent imaging, we initialize $o_c(\mathbf{r})$ with all ones. The pattern, $p_c(\mathbf{r})$, is initialized with the mean of the square root of registered coherent in-focus intensity stack. The pupil function is initialized with a circ function (2D function filled with ones within the defined radius) with the radius defined by the objective NA. In the end, we initialize the scanning position, $\mathbf{r}_{\ell z}$, from the registration of the intensity stacks, $I_{c,\ell z}$, for respective focal planes.

For the coherent imaging reconstruction, we use a total number of $K_c \approx 30$ iterations to converge. We sequentially update $o_c(\mathbf{r})$, $p_c(\mathbf{r})$, $h_c(\mathbf{r})$, and \mathbf{r}_ℓ , ($\ell = 1, \dots, N_{\text{img}}$) for each defocused plane (total number of defocused planes is N_z) per iteration. Unlike for our fluorescence reconstructions, we do not use the extra Nesterov's acceleration step in the QP reconstruction.

Algorithm 3 Coherent imaging reconstruction

Require: $I_{c,\ell z}(\mathbf{r})$, $\mathbf{r}_{\ell z}$, $\ell = 1, \dots, N_{\text{img}}$

- 1: initialize $o_c^{(1,0)}(\mathbf{r})$ with all one values
 - 2: initialize $p_c^{(1,0)}(\mathbf{r}) = \sum_\ell \sqrt{I_{c,\ell,z=0}(\mathbf{r} + \mathbf{r}_{\ell,z=0})} / N_{\text{img}}$
 - 3: initialize $\tilde{h}_c^{(1,0)}(\mathbf{u})$ with all one values within a defined radius set by the objective NA
 - 4: initialize $\mathbf{r}_{\ell z}^{(1)}$ with the scanning position from the registration step
 - 5: **for** $k = 1 : K_c$ **do**
 - 6: Sequential gradient descent
 - 7: **for** $t = 1 : (N_{\text{img}} \cdot N_z)$ **do**
 - 8: $z = z_{\text{mod}(t,2)}$
 - 9: $\ell = \text{mod}(t, N_{\text{img}})$
 - 10: **if** $t < N_{\text{img}} \cdot N_z$ **then**
 - 11: $o_c^{(k,t)}(\mathbf{r}) = o_c^{(k,t-1)}(\mathbf{r}) - \nabla_{o_c} f_{c,\ell z}(o_c^{(k,t-1)}, p_c^{(k,t-1)}, \tilde{h}_c^{(k,t-1)}, \mathbf{r}_{\ell z}^{(k)}) / \max \left(\left| p_c^{(k,t-1)}(\mathbf{r}) \right| \right)^2$
 - 12: $p_c^{(k,t)}(\mathbf{r}) = p_c^{(k,t-1)}(\mathbf{r}) - \nabla_{p_c} f_{c,\ell z}(o_c^{(k,t-1)}, p_c^{(k,t-1)}, \tilde{h}_c^{(k,t-1)}, \mathbf{r}_{\ell z}^{(k)}) / \max \left(\left| o_c^{(k,t-1)}(\mathbf{r}) \right| \right)^2$
 - 13: $\xi(\mathbf{u}) = \mathcal{F}\{o_c^{(k,t-1)}(\mathbf{r}) \cdot \mathcal{C}\{p_c^{(k,t-1)}(\mathbf{r} - \mathbf{r}_\ell)\}\}$
 - 14: $\tilde{h}_c^{(k,t)}(\mathbf{u}) = \tilde{h}_c^{(k,t-1)}(\mathbf{u}) - \nabla_{\tilde{h}_c} f_{c,\ell z}(o_c^{(k,t-1)}, p_c^{(k,t-1)}, \tilde{h}_c^{(k,t-1)}, \mathbf{r}_{\ell z}^{(k)}) \cdot |\xi(\mathbf{u})| / 5[\max(|\xi(\mathbf{u})|) \cdot (|\xi(\mathbf{u})|^2 + \delta)]$, where δ is chosen to be small
 - 15: **else**
 - 16: Do the same update but save to $o_c^{(k+1,0)}(\mathbf{r}), p_c^{(k+1,0)}(\mathbf{r}), \tilde{h}_c^{(k+1,0)}(\mathbf{r})$
 - 17: **end if**
 - 18:
 - 19: Scanning position refinement
 - 20: $x_{\ell z}^{(k+1)} = x_{\ell z}^{(k)} - \beta \nabla_{x_{\ell z}} f_{c,\ell z}(o_c^{(k,t-1)}, p_c^{(k,t-1)}, \tilde{h}_c^{(k,t-1)}, \mathbf{r}_{\ell z}^{(k)})$
 - 21: $y_{\ell z}^{(k+1)} = y_{\ell z}^{(k)} - \beta \nabla_{y_{\ell z}} f_{c,\ell z}(o_c^{(k,t-1)}, p_c^{(k,t-1)}, \tilde{h}_c^{(k,t-1)}, \mathbf{r}_{\ell z}^{(k)})$
 - 22: **end for**
 - 23: **end for**
-

4.3 Experimental results

Figure 4.1 shows our experimental setup. A green laser beam (BeamQ, 532 nm, 200 mW) is collimated through a single lens. The resulting plane wave illuminates the layered Scotch tape (4 layers of 3M 810 Scotch Tape, S-9783), creating a speckle pattern at the sample. The Scotch tape is mounted on a 3-axis piezo-stage (Thorlabs, MAX311D) to enable lateral speckle scanning. The transmitted light from the sample then travels through a $4f$ system formed by the objective lens (OBJ) and a single lens. In order to control the NA of our detection system (necessary for our verification experiment), an extra $4f$ system with an adjustable iris-aperture (AP) in the Fourier space is added. Then, the coherent and fluorescent light are optically separated by a dichroic mirror (DM, Thorlabs, DMLP550R), since they have different wavelengths. The fluorescence is further spectrally filtered (SF) before imaging onto Sensor-F (PCO.edge 5.5). The (much brighter) coherent light is ND-filtered and then split by another beam-splitter before falling on the two defocused coherent sensors, Sensor-C1 and Sensor-C2 (FLIR, BFS-U3-200S6M-C). Sensor-C1 is focused on the sample, while Sensor-C2 is defocused by 0.8 mm.

For our initial verification experiments, we use a $40\times$ objective (Nikon, CFI Achrom 40 \times) with $\text{NA} = 0.65$ as our system’s microscope objective (OBJ). Later high-content experimental demonstrations switch to a $4\times$ objective (Nikon, CFI Plan Achrom 4 \times) with $\text{NA} = 0.1$.

4.3.1 Sample preparation

Results presented in this work targeted super-resolution of $1\ \mu\text{m}$ and $2\ \mu\text{m}$ diameter polystyrene microspheres (Thermofischer) that were fluorescently tagged to emit at a center wavelength of $\lambda_{\text{em}} = 605\ \text{nm}$. Monolayer samples of these microspheres were prepared by placing microsphere dilutions (60 uL stock-solution/500 uL isopropyl alcohol) onto #1.5 coverslips and then allowing to air-dry. High-index oil ($n_m(\lambda) = 1.52$ at $\lambda = 532\ \text{nm}$) was subsequently placed on the coverslip to index-match the microspheres. An adhesive spacer followed by another #1.5 coverslip was placed on top of the original coverslip to assure a uniform sample layer for imaging.

4.3.2 Super-resolution verification

4.3.2.1 Fluorescent imaging verification

We start with a proof-of-concept experiment to verify that our method accurately reconstructs a fluorescent sample at resolutions greater than twice the imaging system’s diffraction-limit. To do so, we use the higher-resolution objective ($40\times$, $\text{NA} 0.65$) and a tunable Fourier-space iris-aperture (AP) that allows us to artificially reduce the system’s NA (NA_{sys}), and therefore, resolution. With the aperture mostly closed (to $\text{NA}_{\text{sys}} = 0.1$), we acquire a low-resolution SIM dataset, which is then used to computationally reconstruct a super-resolved image of the sample with resolution corresponding to an effective $\text{NA} = 0.4$. This recon-

struction is then compared to the widefield image of the sample acquired with the aperture open to $\text{NA}_{\text{sys}} = 0.4$, for validation.

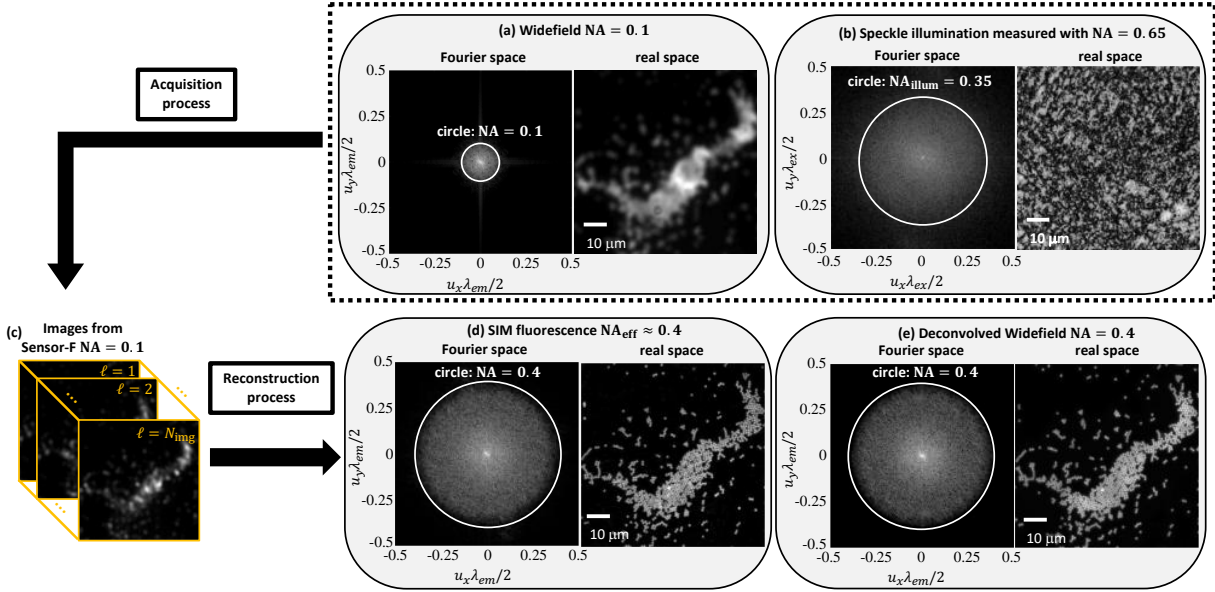


Figure 4.2: Verification of fluorescence super-resolution with $4\times$ resolution gain. Widefield images, for comparison, were acquired at (a) 0.1 NA and (e) 0.4 NA by adjusting the aperture size. (b) The Scotch tape speckle pattern creates much higher spatial frequencies (~ 0.35 NA) than the 0.1 NA detection system can measure. (c) Using the 0.1 NA aperture, we acquire low-resolution fluorescence images for different lateral positions of the Scotch tape. (d) The reconstructed SIM image contains spatial frequencies up to ~ 0.4 NA and is in agreement with (e) the deconvolved widefield image with the system operating at 0.4 NA.

Results are shown in Fig. 4.2, comparing our method against widefield fluorescence images at NAs of 0.1 and 0.4, with no Scotch tape in place. The sample is a monolayer of $1 \mu\text{m}$ diameter microspheres, with center emission wavelength $\lambda_{\text{em}} = 605 \text{ nm}$. At 0.1 NA, the expected resolution is $\lambda_{\text{em}}/2\text{NA} \approx 3.0 \mu\text{m}$ and the microspheres are completely unresolvable. At 0.4 NA, the expected resolution is $\lambda_{\text{em}}/2\text{NA} \approx 0.76 \mu\text{m}$ and the microspheres are well-resolved. With Scotch tape and 0.1 NA, we acquire a set of measurements as we translate the speckle pattern in 267 nm increments on a 26×26 rectangular grid - $N_{\text{img}} = 676$ acquisitions total (details in Sec. 5.4).

Figure 4.2(d) shows the final SR reconstruction of the fluorescent sample in real space, along with the amplitude of its Fourier spectrum. Individual microspheres can be clearly resolved, and results match well with the 0.4 NA deconvolved widefield image (Fig. 4.2(e)). Fourier-space analysis confirms our resolution improvement factor to be $4\times$, which suggests that the Scotch tape produces $\text{NA}_{\text{illum}} \approx 0.3$. To verify, we fully open the aperture and observe that the speckle pattern contains spatial frequencies up to $\text{NA}_{\text{illum}} \approx 0.35$ (Fig. 4.2(b)).

4.3.2.2 Coherent imaging verification

To quantify super-resolution in the coherent imaging channel, we use the low-resolution objective ($4\times$, NA 0.1) to image a USAF1951 resolution chart (Benchmark Technologies). This phase target provides different feature sizes with known phase values, so is a suitable calibration target to quantify both the coherent resolution and the phase sensitivity of our technique.

Results are shown in Fig. 4.3. The coherent intensity image (Fig. 4.3(a)) acquired with 0.1 NA (no tape) has low resolution ($\sim 5.32 \mu\text{m}$), so hardly any features can be resolved. In Fig. 4.3(b), we show the “ground truth” QP distribution at 0.4 NA, as provided by the manufacturer.

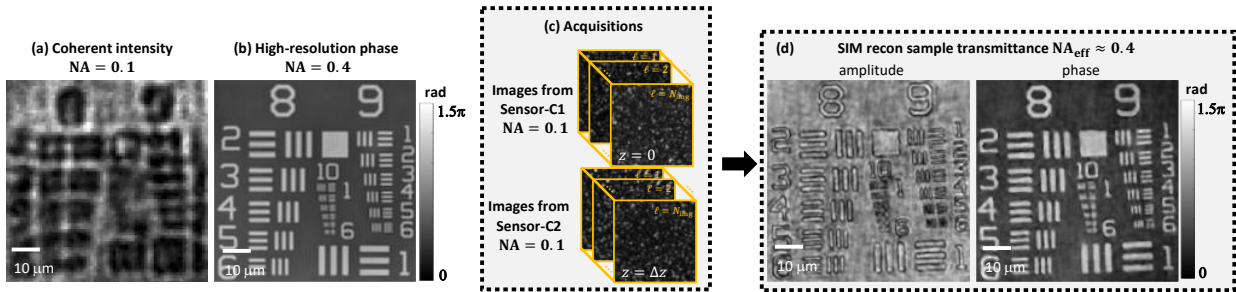


Figure 4.3: Verification of coherent quantitative phase (QP) super-resolution with $4\times$ resolution gain. (a) Low-resolution intensity image and (b) “ground truth” phase at NA=0.4, for comparison. (c) Raw acquisitions of the speckle-illuminated sample intensity from two focus planes, collected with 0.1 NA. (d) Reconstructed SR amplitude and QP, demonstrating $4\times$ resolution gain.

After inserting the Scotch tape, it was translated in 400 nm increments on a 36×36 rectangular grid, giving $N_{\text{img}} = 1296$ total acquisitions (details in Sec. 5.4) at each of the two defocused coherent sensors (Fig. 4.3(c)). Figure 4.3(d,e) shows the SR reconstruction for the amplitude and phase of this sample, resolving features up to group 9 element 5 ($1.23 \mu\text{m}$ separation). Thus, our coherent reconstruction has a $\sim 4\times$ resolution gain compared to the brightfield intensity image.

4.3.3 High-content fluorescent and QP microscopy

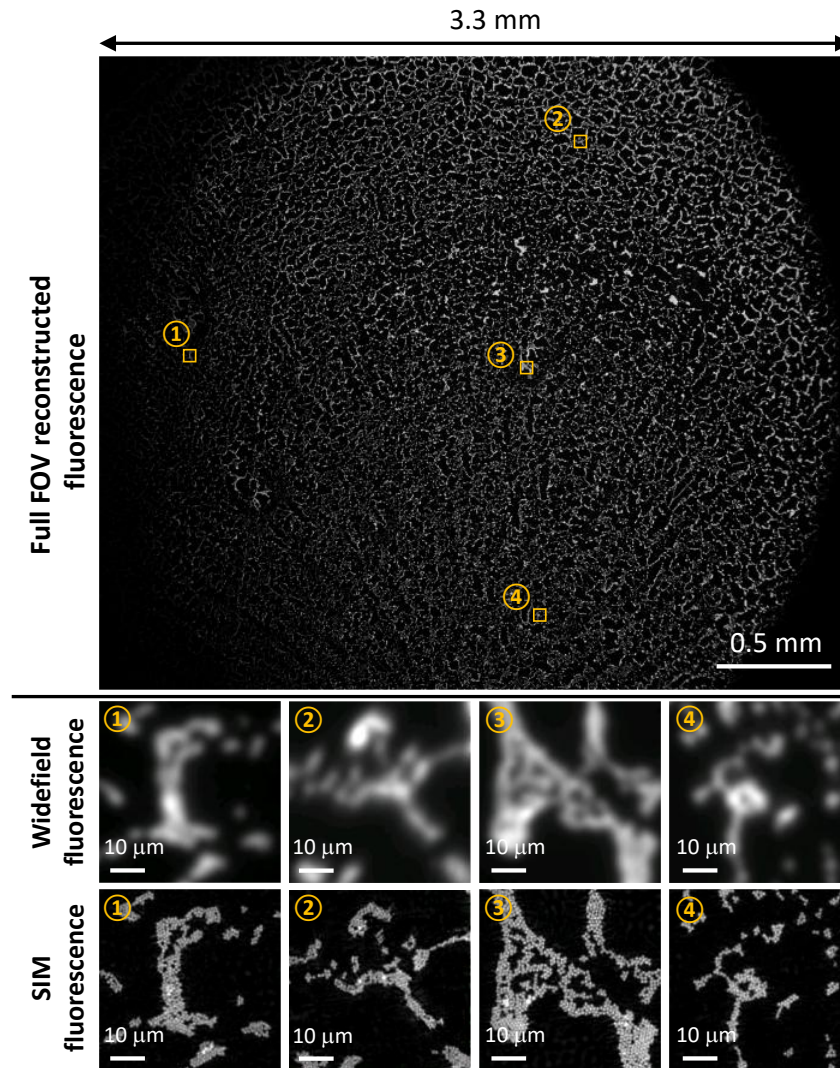


Figure 4.4: Reconstructed super-resolution fluorescence with $4\times$ resolution gain across the full FOV (See Visualization 1). Four zoom-ins of regions-of-interest (ROIs) are compared to their widefield counterparts.

Of course, artificially reducing resolution in order to validate our method required using a moderate-NA objective, which precluded imaging over the large FOVs allowed by low-NA objectives. In this section, we demonstrate high-content fluorescence imaging with the low-resolution, large FOV objective ($4\times$, NA 0.1) to visualize a $2.7\times 3.3\text{ mm}^2$ FOV (see Fig. 4.4(a)). We note that this FOV is more than $100\times$ larger than that allowed by the $40\times$ objective used in the verification experiments, so is suitable for large SBP imaging.

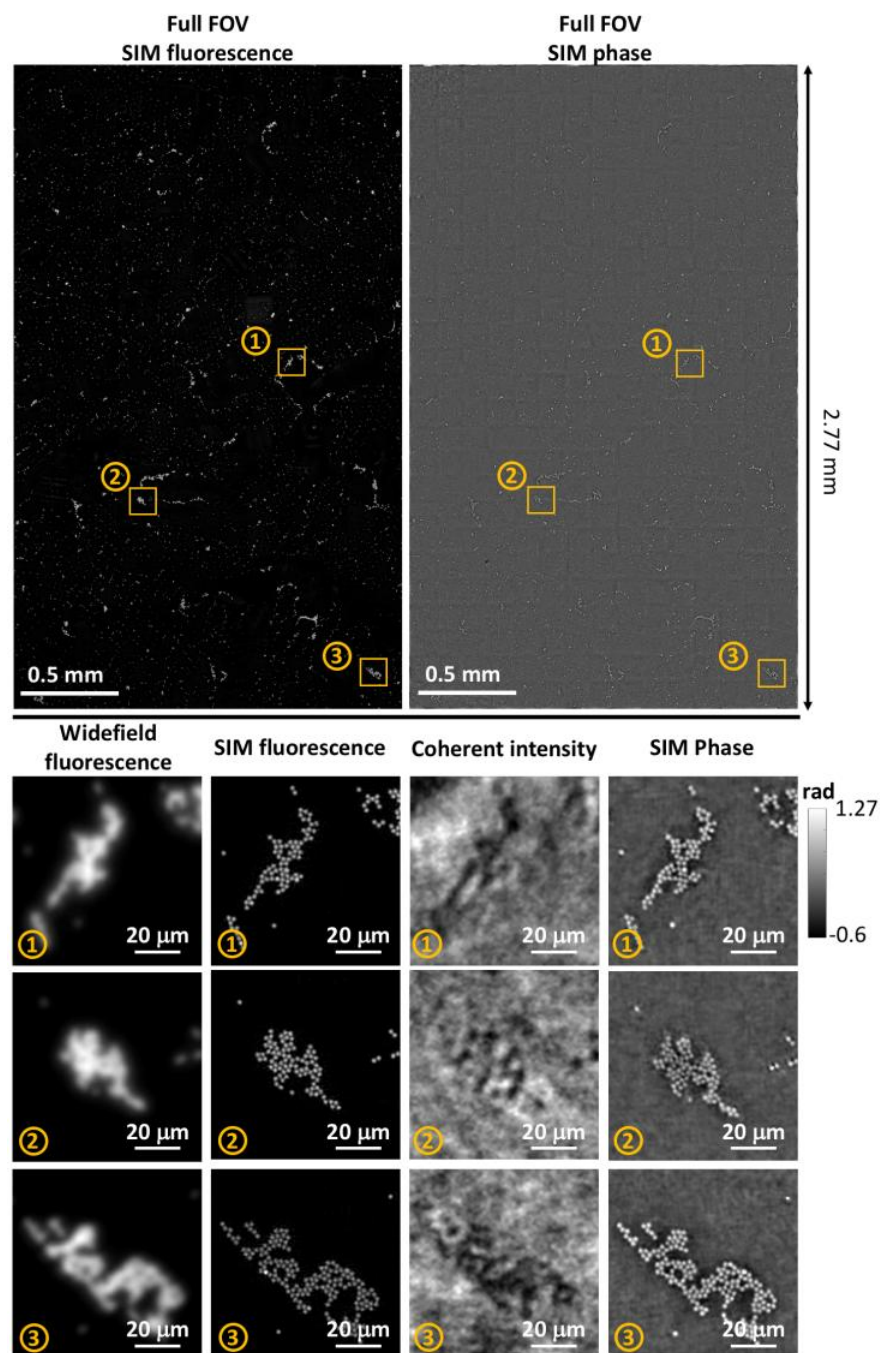


Figure 4.5: Reconstructed multimodal (fluorescence and quantitative phase) high-content imaging (See Visualization 2 and 3). Zoom-ins for three ROIs compare the widefield, super-resolved fluorescence, coherent intensity, and super-resolved phase reconstructions.

Within the imaged FOV for our $1\ \mu\text{m}$ diameter microsphere monolayer sample, we zoom in to four regions-of-interest (ROI), labeled (1), (2), (3), and (4). Widefield fluorescence

imaging cannot resolve individual microspheres, as expected. Using our method, however, gives a factor $4\times$ resolution gain across the whole FOV and enables resolution of individual microspheres. Thus, the SBP of the system, natively ~ 5.3 mega-pixels of content, was increased to ~ 85 mega-pixels, a factor of $4^2 = 16\times$. Though this is still not in the Gigapixel range, this technique is scalable and could reach that range with a higher-SBP objective and sensors.

We next include the QP imaging channel to demonstrate high-content multimodal imaging, as shown in Fig. 4.5. The multimodal FOV is smaller ($2\times 2.7\text{ mm}^2$ FOV) than that presented in Fig. 4.4 because our coherent detection sensors have a lower pixel-count than our fluorescence detection sensor. Figure 4.5 includes zoom-ins of three ROIs to visualize the multimodal SR.

As expected, the widefield fluorescence image and the on-axis coherent intensity image do not allow resolution of individual $2\text{ }\mu\text{m}$ microspheres, since the theoretical resolution for fluorescence imaging is $\lambda_{\text{em}}/2\text{NA}_{\text{sys}} \approx 3\text{ }\mu\text{m}$ and for QP imaging is $\lambda_{\text{ex}}/\text{NA}_{\text{sys}} \approx 5\text{ }\mu\text{m}$. However, our SIM reconstruction with $4\times$ resolution gain enables clear separation of the microspheres in both channels. Our fluorescence and QP reconstructions match well, which is expected since the fluorescent and QP signal originate from identical physical structures in this particular sample.

The full-FOV reconstructions (Fig. 4.4 and 4.5) are obtained by dividing the FOV into small patches, reconstructing each patch, then stitching together the high-content images. Patch-wise reconstruction is computationally favorable because of its low-memory requirement, but also allows us to correct field-dependent aberrations. Since we process each patch separately using our self-calibration algorithm, we solve for each patch’s PSF independently and correct the local aberrations digitally. The reconstruction takes approximately 15 minutes for each channel on a high-end GPU (NVIDIA, TITAN Xp) for a patch with FOV of $110 \times 110\text{ }\mu\text{m}^2$.

4.3.4 Self-calibration analysis

In Sec. 4.2.1.2 and 4.2.2.2, we presented the inverse problem formulation for super-resolution fluorescence and QP. We note that those formulations also included terms to self-calibrate for unknowns in the system’s experimental OTF and the illumination pattern’s scan-position. Here we demonstrate how these calibrations are important for our reconstruction quality.

To demonstrate the improvement in our fluorescence imaging reconstruction due to the self-calibration algorithm, we select a region of interest from the dataset presented in Fig. 4.4. Figure 4.6 shows the comparison of the SR reconstruction with and without self-calibration. The SR reconstruction with no self-calibration contains severe artifacts in reconstructions of both the speckle illumination pattern and the sample’s fluorescent distribution. With OTF correction, dramatic improvements in the fluorescence SR image are evident. OTF correction is especially important when imaging across a large FOV (Fig. 4.4 and 4.5) due to space-varying aberrations.

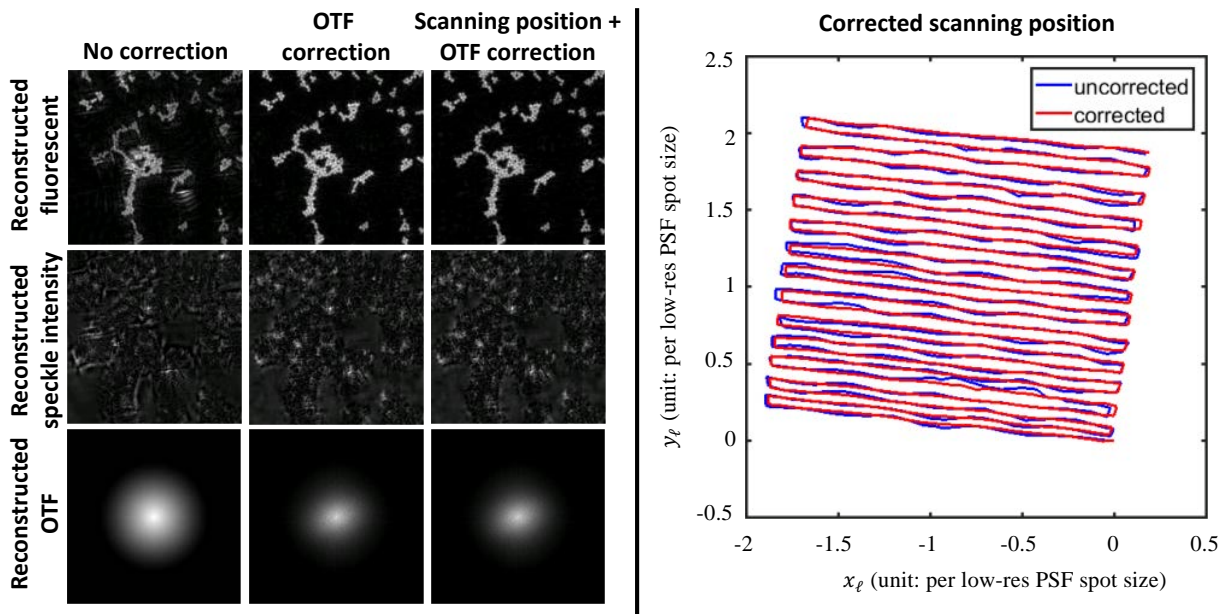


Figure 4.6: Algorithmic self-calibration significantly improves fluorescence super-resolution reconstructions. Here, we compare the reconstructed fluorescence image, speckle intensity, and OTF with no correction, OTF correction, and both OTF correction and scanning position correction. The right panel shows the overlay of the uncorrected and corrected scanning position trajectories.

Further self-calibration to correct for errors in the initial estimate of the illumination pattern’s trajectory enables further refinement of the SR reconstruction. We see that this illumination trajectory demonstrates greater smoothness after undergoing self-calibration. We fully expect that this calibration step to have important ramifications in cases where the physical translation stage is of lower stability or more inaccurate incremental translation.

We also test how the self-calibration affects our phase reconstruction, using the same dataset as in Fig. 4.3. Similar to the conclusion from the fluorescence self-calibration demonstration, pupil correction (coherent OTF) plays an important role in reducing SR reconstruction artifacts as shown in Fig. 4.7. The reconstructed pupil phase suggests that our system aberration is mainly caused by astigmatism. Further refinement of the trajectory of the illumination pattern improves the SR resolution by resolving one more element (group 9 element 6) of the USAF chart. Paying more attention to the uncorrected and corrected illumination trajectory, we find that the self-calibrated trajectory of the illumination pattern tends to align the trajectories from the two coherent cameras. We also notice that the trajectory from the quantitative-phase channels seems to jitter more compared to the fluorescence channel. We hypothesize that this is due to longer exposure time for each fluorescence acquisition, which would average out the jitter.

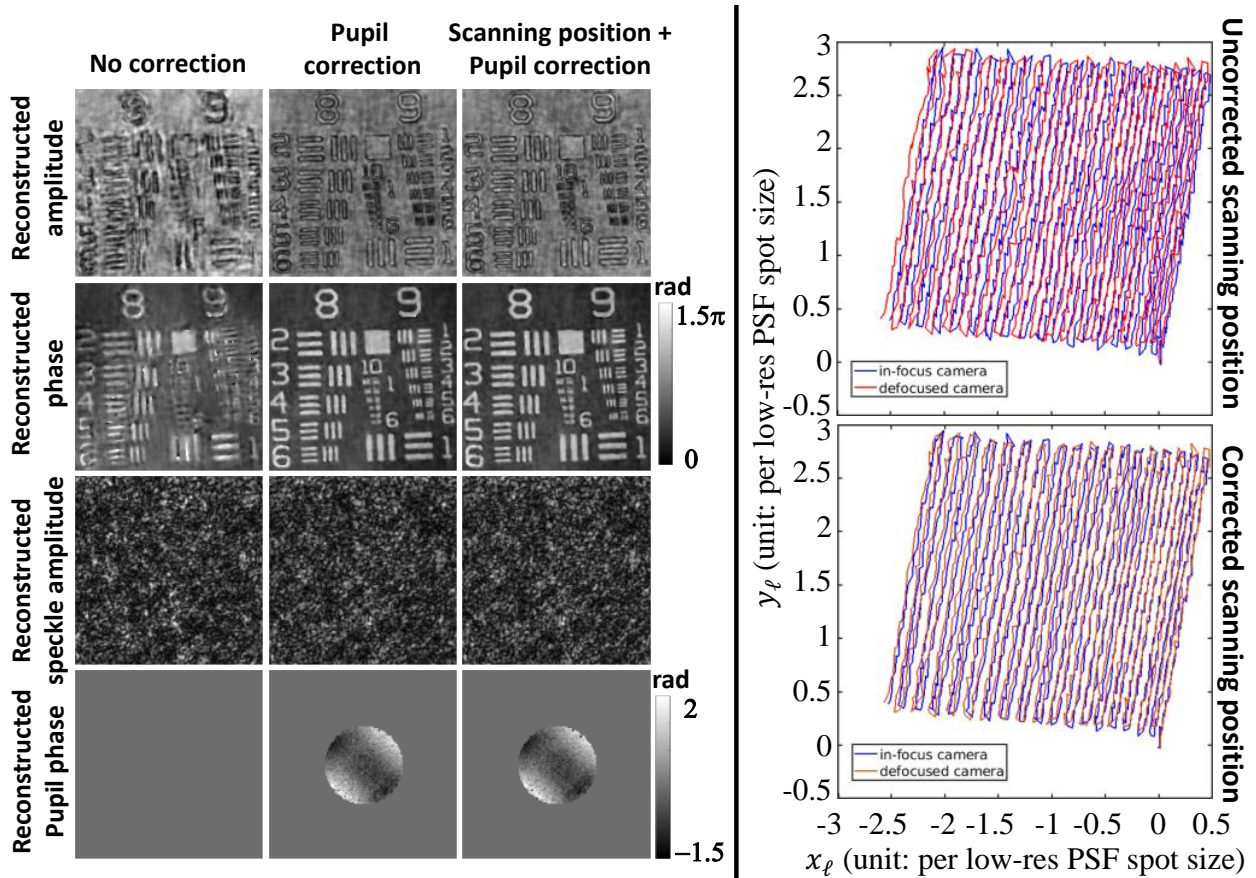


Figure 4.7: Algorithmic self-calibration significantly improves coherent super-resolution reconstructions. We show a comparison of reconstructed amplitude, phase, speckle amplitude, and phase of the pupil function with no correction, pupil correction, and both pupil correction and scanning position correction. The right panel shows the overlay of scanning position trajectory for the in-focus and defocused cameras before and after correction.

4.4 Discussion

Unlike many existing high-content imaging techniques, one benefit of our method is its easy compatibility for simultaneous QP and fluorescence imaging. This arises from SIM’s unique ability to multiplex both coherent and incoherent signals into the system aperture [13]. Furthermore, existing high-content fluorescence imaging techniques that use micro-lens arrays [23–28] are resolution-limited by the physical size of the lenslets, which typically have $NA_{\text{illum}} < 0.3$. Recent work [29] has introduced a framework in which gratings with sub-diffraction slits allow sub-micron resolution across large FOVs - however, this work is heavily limited by SNR, due to the primarily opaque grating, as well as tight required axial alignment. Though the Scotch tape used in our proof-of-concept prototype also induced illumination

angles within a similar range as micro-lens arrays ($\text{NA}_{\text{illum}} \approx 0.35$), we could in future use a stronger scattering media to achieve $\text{NA}_{\text{illum}} \approx 1.0$, enabling further SR and thus larger SBP.

The main drawback of our technique is that we use around ~ 1200 translations of the Scotch tape for each reconstruction, which results in long acquisition times (~ 180 seconds for shifting, pausing, and capturing) and higher photon requirements. Heuristically, for both fluorescence and QP imaging, we found that a sufficiently large scanning range (larger than ~ 2 low-NA diffraction limited spot sizes) and finer scan steps (smaller than the targeted resolution) can reduce distortions in the reconstruction. Tuning such parameters to minimize the number of acquisitions without degrading reconstruction quality is thus an important subject for future endeavors.

4.5 Conclusion

We have presented a large-FOV multimodal SIM fluorescence and QP imaging technique. We use Scotch tape to efficiently generate high-resolution features over a large FOV, which can then be measured with both fluorescent and coherent contrast using a low-NA objective. A computational optimization-based self-calibration algorithm corrected for experimental uncertainties (scanning-position, aberrations, and random speckle pattern) and enabled super-resolution fluorescence and quantitative phase reconstruction with factor $4\times$ resolution gain.

Chapter 5

Computational speckle structured illumination 3D phase and fluorescence microscopy

In Chapter 4, we implemented structured illumination microscopy (SIM) using speckle illumination from simple Scotch tape as a solution for 2D multimodal high-content microscopy. In this work, we extend the method to 3D by leveraging the fact that the speckle illumination is in fact a 3D structured pattern. We use both a coherent and an incoherent imaging model to develop algorithms for joint retrieval of the 3D super-resolved fluorescent and complex-field distributions of the sample. Our reconstructed images resolve features beyond the physical diffraction-limit set by the system’s objective ($> 2\times$ lateral resolution improvement for the coherent channel) and demonstrate 3D multimodal imaging with $\sim 0.6 \times 0.6 \times 6 \mu\text{m}^3$ resolution over a volume of $\sim 314 \times 500 \times 24 \mu\text{m}^3$.

5.1 Introduction

High-content optical microscopy is a driving force for large-scale biological study in fields such as drug discovery and systems biology. With fast imaging speeds over large fields-of-view (FOV) and high spatial resolutions [31–37, 133], one can visualize rare cell phenotypes and dynamics. The traditional solution for 2D high-content microscopy is to mechanically scan samples through the limited FOV of a high-NA (*i.e.* high resolution) imaging objective and then digitally stitch the images together. However, this scheme is limited in imaging speed due to the large-distance translations of the sample as well as the need for auto-refocusing at each translation position [134]. These issues are further compounded when extending this high-content imaging strategy to 3D.

Recently, computational imaging has demonstrated efficient strategies for high-content 2D microscopy. In contrast with slide scanning, these strategies often employ a low-NA imaging objective to acquire low-resolution (large-FOV) measurements. Developments in

synthetic aperture [2, 11, 41] and super-resolution (SR) [3, 5–8, 58] microscopies have demonstrated that a series of such measurements, if taken under varying sample illuminations, can be digitally combined into a single high-resolution large-FOV image. This eliminates the requirement for large-distance mechanical scanning in high-content imaging, which results in faster acquisition and more cost-effective optical setups, while also relaxing the sample’s auto-focusing requirements due to the low-NA objective’s longer depth-of-field (DOF) [17–30, 79, 148, 149]. Examples of such approaches include lensless microscopy [17–19] and Fourier ptychography [20–22, 79, 148, 149] for coherent absorption and quantitative phase imaging. For incoherent fluorescent imaging, micro-lenslet arrays [24, 25, 27, 28], Talbot plane scanning [23, 26, 29], diffuse media [30], or meta-surfaces [150] have also been demonstrated. Among these examples, 3D high-content imaging capability has only been demonstrated in the coherent imaging context (quantitative phase and absorption) by Fourier ptychography [79, 149].

Our previous work demonstrated multimodal coherent (quantitative phase) and incoherent (fluorescence) imaging for high-content 2D microscopy [151]. Multimodal imaging is important for biological studies requiring cross-correlative analysis [13, 16, 140, 141, 152]. Structured illumination microscopy (SIM) [2, 5, 6, 60] with speckle illumination [30, 51–54, 56, 57, 113, 114, 130] has been used to encode 2D SR quantitative phase and fluorescence. However, because propagating speckle contains 3D features, it also encodes 3D information. Considering speckle patterns as random interference of multiple angled plane waves, the scattered light from interactions with the sample carries 3D phase (coherent) information, similar to the case of non-random angled illumination in diffraction tomography [9, 153–155] and 3D Fourier ptychography [79, 149]. Simultaneously, the fluorescent (incoherent) light excited by the 3D speckle pattern encodes 3D SR fluorescence information as in the case of 3D SIM [129]. Combining these, we propose a method for 3D SR quantitative phase and fluorescence microscopy using speckle illumination.

Experimentally, we position a Scotch tape patterning element just before the sample, mounted on a translation stage to generate a translating speckle field that illuminates the sample (Fig. 5.1). Because the speckle grain size is smaller than the PSF of the low-NA imaging objective (which provides large-FOV), the coherent scattered light from the speckle-sample interaction encodes 3D SR quantitative phase information. In addition to lateral scanning of the Scotch tape, axial sample scanning is necessary to efficiently capture 3D SR fluorescence information. Nonlinear computational methods based on the 3D coherent beam propagation model [79, 156] and the 3D incoherent imaging model [129] were formulated to reconstruct the 3D speckle field and imaging system aberrations, which are subsequently used to reconstruct the sample’s 3D SR quantitative phase and fluorescence distributions. Since the Scotch tape is directly before the sample, the illumination NA is not limited by the objective lens, allowing for $> 2\times$ lateral resolution gain across the entire FOV. This framework enables us to achieve 3D imaging at sub-micron lateral resolution and micron axial resolution across a half-millimeter FOV.

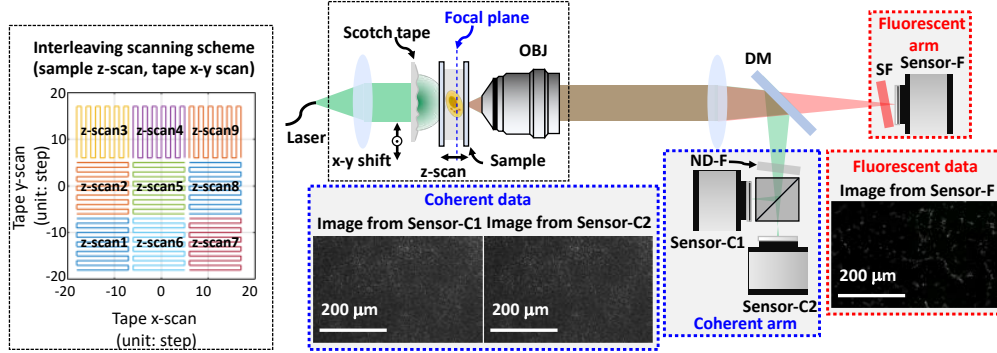


Figure 5.1: 3D multimodal structured illumination microscopy (SIM) with laterally translating Scotch tape as the patterning element. The coherent arm (Sensor-C1 and Sensor-C2) simultaneously captures images with different defocus at the laser illumination wavelength ($\lambda_{\text{ex}} = 532 \text{ nm}$), used for both 3D phase retrieval and speckle trajectory calibration. The incoherent (fluorescence) arm (Sensor-F) captures low-resolution raw fluorescence acquisitions at the emission wavelength ($\lambda_{\text{em}} = 605 \text{ nm}$) for 3D fluorescence super-resolution reconstruction. OBJ: objective, DM: dichroic mirror, SF: spectral filter, ND-F: neutral-density filter.

5.2 Theory

We start by introducing the concept of 3D coherent and incoherent transfer functions (TFs), using the Born (weak scattering) assumption [153], to analyze the information encoding process. We then lay out our 3D coherent and incoherent imaging models and derive the corresponding inverse problems to extract SR quantitative phase and fluorescence from the measurements.

Previous work derived linear space-invariant relationships between raw measurements and 3D coherent scattering and incoherent fluorescence [129, 153, 157, 158]. These relationships enable us to define TFs for the coherent and incoherent imaging processes. The supports of these TFs in 3D Fourier space determine how much spatial frequency content of the sample can be passed through the system (*i.e.* the 3D diffraction-limited resolution).

In a coherent imaging system with on-axis plane-wave illumination, the TF describes the relationship between the sample’s scattering potential and the measured 3D scattered field, and takes the shape of a spherical cap in 3D Fourier space (Fig. 5.2(a)). In an incoherent imaging system, the TF is the autocorrelation of the coherent system’s TF [158], relating the sample’s fluorescence distribution to the 3D measured intensity. It takes the shape of a torus (Fig. 5.2(b)). The spatial frequency bandwidth of these TFs are summarized in Table 5.1, where the lateral resolution of the system is proportional to the lateral bandwidth of the TF. The incoherent TF has $2\times$ greater bandwidth than the coherent TF and axial bandwidth generally depends on the lateral spatial frequency, so axial resolution is specified in terms of the best-case. Note that the axial bandwidth of the coherent TF is zero, which means there is zero axial resolution for coherent imaging; hence the poor depth sectioning ability in 3D

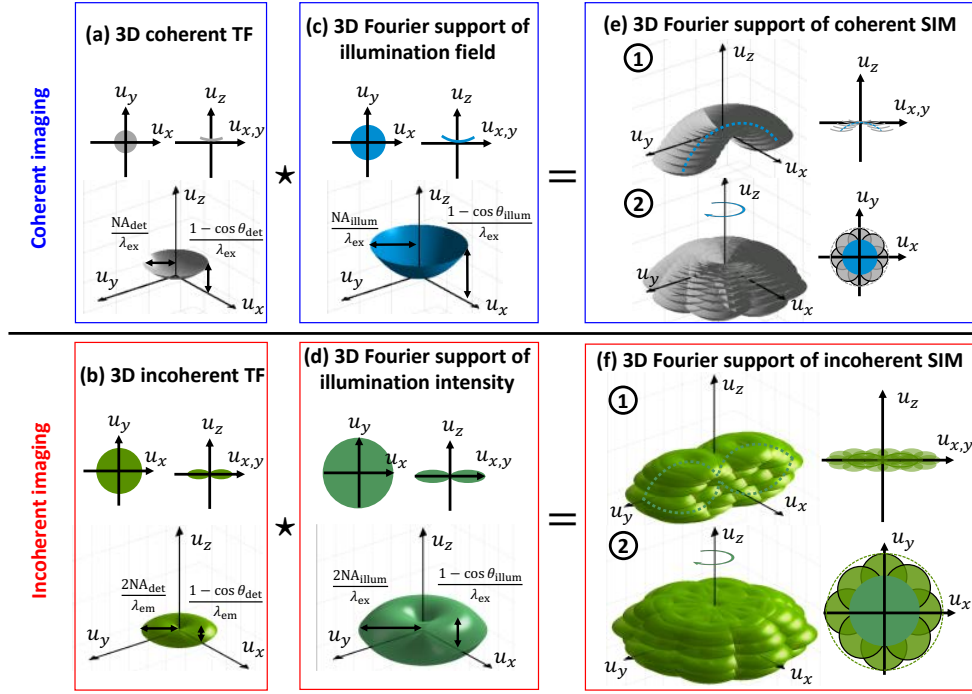


Figure 5.2: 3D coherent and incoherent transfer function (TF) analysis of the SIM imaging process. The 3D (a) coherent and (b) incoherent TFs of the detection system are autocorrelated with the 3D Fourier support of the (c) illumination speckle field and (d) illumination intensity, respectively, resulting in the effective Fourier support of 3D (e) coherent and (f) incoherent SIM. In (e) and (f), we display decomposition of the auto-correlation in two steps: ① tracing the illumination support in one orientation and ② replicating this trace in azimuthal direction.

holographic imaging [16, 155, 159].

SIM enhances resolution by creating beat patterns. When a 3D structured pattern modulates the sample, the sample's sub-diffraction features create lower-frequency beat patterns which can be directly visualized and used to reconstruct a SR image of the sample via post-processing [6, 129]. This process is generally applicable to both coherent and incoherent imaging [13, 16, 140, 141], enabling 3D SR multimodal imaging. Mathematically, a modulation between the sample contrast and the illumination pattern in real space can be interpreted as a convolution in Fourier space. This convolution result is then passed through the 3D TF defined in Fig. 5.2(a,b). The effective support of information going into the measurements can be estimated by conducting cross-correlations between the 3D TFs and the Fourier content of the illumination patterns, as shown in Fig. 5.2(c,e) and 5.2(d,f) for coherent and incoherent systems, respectively. The lateral and axial spatial frequency bandwidth of both illumination and 3D SIM Fourier supports for coherent and incoherent imaging are summarized in Table 5.1. Assuming approximately equal excitation and

Table 5.1: Summary of spatial frequency bandwidths

	Lateral bandwidth	Axial bandwidth
Coherent TF	$\frac{2NA_{\text{det}}}{\lambda_{\text{ex}}}$	0
Incoherent TF	$\frac{4NA_{\text{det}}}{\lambda_{\text{em}}}$	$\frac{2(1-\cos\theta_{\text{det}})}{\lambda_{\text{em}}}$
Illum. field	$\frac{2NA_{\text{illum}}}{\lambda_{\text{ex}}}$	0
Illum. intensity	$\frac{4NA_{\text{illum}}}{\lambda_{\text{ex}}}$	$\frac{2(1-\cos\theta_{\text{illum}})}{\lambda_{\text{ex}}}$
3D coherent SIM	$\frac{2NA_{\text{det}}+2NA_{\text{illum}}}{\lambda_{\text{ex}}}$	$\frac{1-\cos\theta_{\text{det}}}{\lambda_{\text{ex}}} + \frac{1-\cos\theta_{\text{illum}}}{\lambda_{\text{ex}}}$
3D incoherent SIM	$\frac{4NA_{\text{det}}}{\lambda_{\text{em}}} + \frac{4NA_{\text{illum}}}{\lambda_{\text{ex}}}$	$2\left(\frac{1-\cos\theta_{\text{det}}}{\lambda_{\text{em}}} + \frac{1-\cos\theta_{\text{illum}}}{\lambda_{\text{ex}}}\right)$

$NA_{\text{det}}, NA_{\text{illum}}$: the numerical aperture of the detection and illumination system,
 $\theta_{\text{det}}, \theta_{\text{illum}}$: the maximal detectable and illuminating half angle of light,
 $\lambda_{\text{ex}}, \lambda_{\text{em}}$: the wavelength of the excitation and emission light.

emission wavelengths, the achievable lateral resolution gain of 3D SIM (ratio between lateral bandwidths of 3D SIM and 3D TF) is $(NA_{\text{det}} + NA_{\text{illum}})/NA_{\text{det}}$ for both coherent and incoherent imaging. Axially, coherent SIM builds up the spatial frequency bandwidth in the axial direction, and incoherent SIM can achieve axial resolution gain with a factor of $(2 - \cos\theta_{\text{det}} - \cos\theta_{\text{illum}})/(1 - \cos\theta_{\text{det}})$.

In this work, the Scotch tape allows us to create a high-resolution speckle illumination such that $NA_{\text{illu}} > NA_{\text{det}}$, enabling $> 2\times$ lateral resolution gain [151]. From the TF analysis, we also see that information beyond diffraction-limit in the axial dimension is obtainable. The next sections outline our computational scheme for 3D SR phase and fluorescence reconstruction. Our algorithm jointly estimates the illumination speckle field, system pupil function (aberrations), the sample's 3D transmittance function, and the sample's 3D fluorescence distribution.

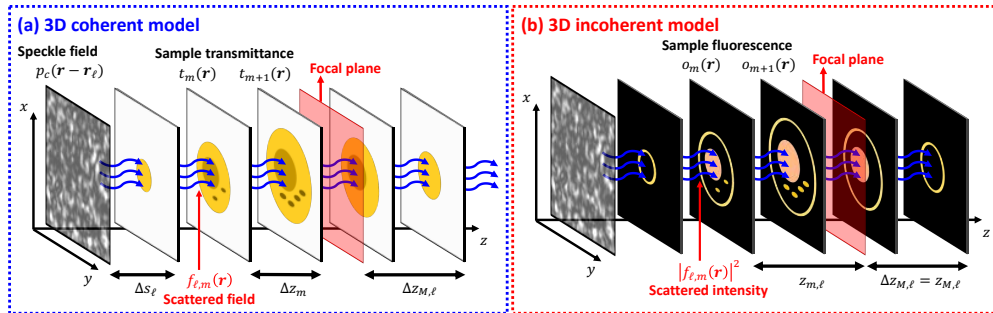


Figure 5.3: 3D multi-slice model: (a) coherent and (b) incoherent imaging models for interaction between the sample and the speckle field.

5.2.1 3D super-resolution phase imaging

We adopt a multi-slice coherent scattering model to describe the 3D multiple-scattering process [79, 156] and solve for 3D SR quantitative phase. Our system captures intensity at two focus planes, z_{c1} and z_{c2} , for every speckle scanned point [151]. With these measurements and the multi-slice model, we are able to reconstruct the sample's 3D SR complex-field and the scattered field inside the 3D sample, which is used in the fluorescence inverse problem.

5.2.1.1 Forward model for 3D coherent imaging

Figure 5.3(a) illustrates the 3D multi-slice coherent imaging model. Plane-wave illumination of the Scotch tape, positioned at the ℓ -th scanned point, \mathbf{r}_ℓ , creates speckle field $p_c(\mathbf{r} - \mathbf{r}_\ell)$, where $\mathbf{r} = (x, y)$ is the lateral spatial coordinate. This speckle field propagates a distance Δs_ℓ to the sample. The field interacting with the first layer of the sample is described as:

$$f_{\ell,1}(\mathbf{r}) = \mathcal{C}\{p_c(\mathbf{r} - \mathbf{r}_\ell) \otimes h_{\Delta s_\ell, \lambda_{\text{ex}}}(\mathbf{r})\}, \quad (5.1)$$

where $h_{z,\lambda}(\mathbf{r}) = \mathcal{F}^{-1} \left\{ \exp \left(i2\pi z \sqrt{1/\lambda^2 - \|\mathbf{u}\|_2^2} \right) \right\}$ is the angular spectrum propagation kernel [38], $\mathbf{u} = (u_x, u_y)$ is the spatial frequency coordinate, and $\mathcal{C}\{\cdot\}$ is a cropping operator that selects the part of the speckle field that illuminates the sample. To model scattering and propagation inside the sample, the multi-slice model approximates the 3D sample as multiple slices of complex transmittance function, $t_m(\mathbf{r})$ ($m = 1, \dots, M$), where m is the slice index number. As the field propagates through each slice, it first multiplies with the 2D transmittance function of the sample at that slice, then propagates to the next slice. The spacing between slices is modeled as uniform media of thickness Δz_m . Hence, at each layer we have:

$$\begin{aligned} g_{\ell,m}(\mathbf{r}) &= f_{\ell,m}(\mathbf{r}) \cdot t_m(\mathbf{r}), \quad m = 1, \dots, M, \\ f_{\ell,m+1}(\mathbf{r}) &= g_{\ell,m}(\mathbf{r}) \otimes h_{\Delta z_m, \lambda_{\text{ex}}}(\mathbf{r}), \quad m = 1, \dots, M - 1. \end{aligned} \quad (5.2)$$

After passing through all the slices, the output scattered field, $g_{\ell,M}(\mathbf{r})$, propagates to the focal plane to form $G_\ell(\mathbf{r}) = g_{\ell,M}(\mathbf{r}) \otimes h_{\Delta z_{M,\ell}, \lambda_{\text{ex}}}(\mathbf{r})$ and gets imaged onto the sensor (with defocus z), forming our measured intensity:

$$I_{c,\ell z}(\mathbf{r}) = |G_\ell(\mathbf{r}) \otimes h_c(\mathbf{r}) \otimes h_{z,\lambda_{\text{ex}}}(\mathbf{r})|^2, \quad \ell = 1, \dots, N_{\text{img}}, \quad z = z_{c1}, z_{c2}, \quad (5.3)$$

where $h_c(\mathbf{r})$ is the system's coherent point spread function (PSF). The measured intensity subscripts c and ℓ denote indices for the coherent imaging channel and acquisition number, respectively. N_{img} is the total number of translations of the Scotch tape. Note that all the spacing distances, Δz_m , are independent of the axial scanned position, Δs_ℓ , except for the distance to the focal plane, which is $\Delta z_{M,\ell} = \Delta s_\ell + z_0$, where z_0 is the distance from the last layer of the sample to the focal plane (before axial scanning). As the sample is scanned, we account for this shift by propagating extra distance back to the focal plane.

5.2.1.2 Inverse problem for 3D coherent imaging

We take the intensity measurements from both coherent cameras, $\{I_{c,\ell z}(\mathbf{r}) \mid z = z_{c1}, z_{c2}\}$, and the scanning trajectory, \mathbf{r}_ℓ (calculated via standard rigid-body 2D registration [147, 151]), as inputs to jointly estimate the sample's 3D SR transmittance function, $t_1(\mathbf{r}), \dots, t_M(\mathbf{r})$, as well as the illumination complex-field, $p_c(\mathbf{r})$, and the system's coherent PSF, $h_c(\mathbf{r})$, including aberrations.

Based on the forward model in the previous section, we formulate the inverse problem as:

$$\begin{aligned} \underset{t_1, \dots, t_M, p_c, h_c}{\text{minimize}} \quad & e_c(t_1, \dots, t_M, p_c, h_c) = \sum_{\ell, z} e_{c,\ell z}(t_1, \dots, t_M, p_c, h_c) \\ \text{where} \quad & e_{c,\ell z}(t_1, \dots, t_M, p_c, h_c) = \sum_{\mathbf{r}} \left| \sqrt{I_{c,\ell z}(\mathbf{r})} - |G_\ell(\mathbf{r}) \otimes h_c(\mathbf{r}) \otimes h_{z,\lambda_{\text{ex}}}(\mathbf{r})| \right|^2. \end{aligned} \quad (5.4)$$

Here we adopt an amplitude-based cost function, e_c , which minimizes the difference between the measured and estimated coherent amplitude in the presence of noise [144]. In order to solve this optimization problem, we use a sequential gradient descent algorithm [144, 145]. The gradient based on each single measurement is calculated and used to update the sample's transmittance function, illumination speckle field, and coherent PSF. A whole iteration of variable updates is complete after running through all the measurements. In Sec. 5.2.3, we provide a detailed derivation of the gradients and in Sec. 5.2.4 we lay out our reconstruction algorithm.

5.2.2 3D super-resolution fluorescence imaging

Reconstruction of 3D SR images for the fluorescence channel involves an incoherent multi-slice forward model (Fig. 5.3(b)) and a joint inverse problem solver. The coherent result provides a good starting estimate of the 3D speckle intensity throughout the sample, which, together with the fluorescent channel's raw data, is used to reconstruct the sample's 3D SR fluorescence distribution and the system's aberrations at the emission wavelength, λ_{em} .

5.2.2.1 Forward model for 3D fluorescence imaging

The 3D fluorescence distribution is also modeled by multiple slices of 2D distributions, $o_m(\mathbf{r})$ ($m = 1, \dots, M$), as shown in Fig. 5.3(b). Each layer is illuminated by the m -th layer's excitation intensity, $|f_{\ell,m}(\mathbf{r})|^2$, for Scotch tape position \mathbf{r}_ℓ . The excited fluorescent light is mapped onto the sensor through 2D convolutions with the incoherent PSF at different defocus distances, $z_{m,\ell}$. The sum of contributions from different layers form the measured fluorescence intensity:

$$I_{f,\ell}(\mathbf{r}) = \sum_{m=1}^M [o_m(\mathbf{r}) \cdot |f_{\ell,m}(\mathbf{r})|^2] \otimes |h_{f,z_{m,\ell}}(\mathbf{r})|^2, \quad \ell = 1, \dots, N_{\text{img}}, \quad (5.5)$$

where $h_{f,z_{m,\ell}}(\mathbf{r})$ is the coherent PSF at defocus distance $z_{m,\ell}$, which could be further decomposed into $h_{f,z_{m,\ell}}(\mathbf{r}) = h_f(\mathbf{r}) \otimes h_{z_{m,\ell},\lambda_{\text{em}}}(\mathbf{r})$, where $h_f(\mathbf{r})$ is the in-focus system's coherent PSF at λ_{em} . The incoherent PSF is the intensity of the coherent PSF at λ_{em} . The subscript f denotes the fluorescence channel and the defocus distance, $z_{m,\ell}$, depends on the axial scanning position, Δs_ℓ .

5.2.2.2 Inverse problem for 3D fluorescence imaging

The fluorescence inverse problem takes as input the raw fluorescence intensity measurements, $I_{f,\ell}(\mathbf{r})$, the registered scanning trajectory, \mathbf{r}_ℓ , and the 3D estimates from the coherent model, in order to estimate the sample's 3D SR fluorescence distribution and aberrations at the emission wavelength. We also refine the speckle field estimate using the fluorescence measurements.

Based on the incoherent forward model, our 3D SR fluorescence inverse problem is:

$$\begin{aligned} \underset{o_1, \dots, o_M, p_c, h_f}{\text{minimize}} \quad & e_f(o_1, \dots, o_M, p_c, h_f) = \sum_{\ell} e_{f,\ell}(o_1, \dots, o_M, p_c, h_f) \\ \text{where} \quad & e_{f,\ell}(o_1, \dots, o_M, p_c, h_f) = \sum_{\mathbf{r}} \left| I_{f,\ell}(\mathbf{r}) - \sum_{m=1}^M [o_m(\mathbf{r}) \cdot |f_{\ell,m}(\mathbf{r})|^2] \otimes |h_{f,z_{m,\ell}}(\mathbf{r})|^2 \right|^2 \end{aligned} \quad (5.6)$$

where e_f is the cost function. Similar to the coherent inverse problem, we adopt a sequential gradient descent algorithm for estimation of each unknown variable. The detailed derivation of gradients and algorithm implementation are summarized in Sec. 5.2.3 and 5.2.4, respectively.

5.2.3 Solving the inverse problem

5.2.3.1 Vectorial notation

In order to derive the gradient to solve for multivariate optimization problem in Eq. (5.4) and (5.6), it is more convenient to represent our 3D coherent and fluorescent model in the linear algebra vectorial notation in the following sections.

According to Eq. (5.1) and (5.2), we are able to re-express the multi-slice scattering model using the vectorial formulation into

$$\begin{aligned} \mathbf{f}_{\ell,1} &= \mathbf{H}_{\Delta s_\ell, \lambda_{\text{ex}}} \mathbf{Q} \mathbf{S}_\ell \mathbf{p}_c, \\ \mathbf{g}_{\ell,m} &= \text{diag}(\mathbf{f}_{\ell,m}) \mathbf{t}_m, \quad m = 1, \dots, M, \\ \mathbf{f}_{\ell,m+1} &= \mathbf{H}_{\Delta z_m, \lambda_{\text{ex}}} \mathbf{g}_{\ell,m}, \quad m = 1, \dots, M-1, \\ \mathbf{G}_\ell &= \mathbf{H}_{\Delta z_M, \lambda_{\text{ex}}} \mathbf{g}_{\ell,M}, \end{aligned} \quad (5.7)$$

where the boldface symbols are the vectorial representation of the 2D variables in non-boldface form in the original model except for the cropping operator \mathbf{Q} , the shifting operator

\mathbf{S}_ℓ that shifts the speckle pattern with \mathbf{r}_ℓ amount, and the defocus convolution operation expressed as

$$\mathbf{H}_{z,\lambda} = \mathbf{F}^{-1} \text{diag}(\tilde{\mathbf{h}}_{z,\lambda}) \mathbf{F}, \quad (5.8)$$

where \mathbf{F} and \mathbf{F}^{-1} are Fourier and inverse Fourier transform operator, respectively, $\tilde{\mathbf{h}}_{z,\lambda}$ is the vectorized coherent TF for propagation distance z and wavelength λ . With all these equations defined in vectorial form, we rewrite our coherent and fluorescence intensity as

$$\begin{aligned} \mathbf{I}_{c,\ell z} &= |\mathbf{H}_c \mathbf{H}_{z,\lambda_{\text{ex}}} \mathbf{G}_\ell|^2, \\ \mathbf{I}_{f,\ell} &= \sum_{m=1}^M \mathbf{K}_{z_{m,\ell}} \text{diag}(|\mathbf{f}_{\ell,m}|^2) \mathbf{o}_m, \end{aligned} \quad (5.9)$$

where \mathbf{H}_c is also a convolution operation as expressed in Eq. (5.8) with the TF vector, $\tilde{\mathbf{h}}_{z,\lambda}$, replaced by the pupil vector $\tilde{\mathbf{h}}_c$, and

$$\mathbf{K}_{z_{m,\ell}} = \mathbf{F}^{-1} \text{diag} \left(\mathbf{F} |\mathbf{F}^{-1} \text{diag}(\tilde{\mathbf{h}}_{z_{m,\ell},\lambda_{\text{em}}}) \tilde{\mathbf{h}}_f|^2 \right) \mathbf{F} \quad (5.10)$$

is the convolution operation with the incoherent TF at $z_{m,\ell}$.

Next we use this vectorial model to represent the coherent and fluorescent cost functions for a single intensity measurement as

$$\begin{aligned} e_{c,\ell z}(\mathbf{t}_1, \dots, \mathbf{t}_M, \mathbf{p}_c, \tilde{\mathbf{h}}_c) &= \mathbf{e}_{c,\ell z}^T \mathbf{e}_{c,\ell z} = \left\| \sqrt{\mathbf{I}_{c,\ell z}} - |\mathbf{H}_c \mathbf{H}_{z,\lambda_{\text{ex}}} \mathbf{G}_\ell| \right\|_2^2, \\ e_{f,\ell}(\mathbf{o}_1, \dots, \mathbf{o}_M, \mathbf{p}_c, \tilde{\mathbf{h}}_f) &= \mathbf{e}_{f,\ell}^T \mathbf{e}_{f,\ell} = \left\| \mathbf{I}_{f,\ell} - \sum_{m=1}^M \mathbf{K}_{z_{m,\ell}} \text{diag}(|\mathbf{f}_{\ell,m}|^2) \mathbf{o}_m \right\|_2^2, \end{aligned} \quad (5.11)$$

where $\mathbf{e}_{c,\ell z} = \sqrt{\mathbf{I}_{c,\ell z}} - |\mathbf{H}_c \mathbf{H}_{z,\lambda_{\text{ex}}} \mathbf{G}_\ell|$ and $\mathbf{e}_{f,\ell} = \mathbf{I}_{f,\ell} - \sum_{m=1}^M \mathbf{K}_{z_{m,\ell}} \text{diag}(|\mathbf{f}_{\ell,m}|^2) \mathbf{o}_m$ are the coherent and fluorescent cost vectors, respectively.

5.2.3.2 Gradient derivation

The following derivation is based on CR calculus and is similar to the derivation introduced by our previous work [144, 151].

Gradient derivation for 3D coherent imaging

To optimize Eq. (5.4) for $\mathbf{t}_1, \dots, \mathbf{t}_M, \mathbf{p}_c, \tilde{\mathbf{h}}_c$, we need to take the derivative of the coherent cost function with respect to them. We first express the gradients of all the transmittance function vectors, $\mathbf{t}_1, \dots, \mathbf{t}_M$ as

$$\nabla_{\mathbf{t}_m} e_{c,\ell z} = \left(\frac{\partial e_{c,\ell z}}{\partial \mathbf{g}_{\ell,m}} \cdot \frac{\partial \mathbf{g}_{\ell,m}}{\partial \mathbf{t}_m} \right)^\dagger = \text{diag}(\overline{\mathbf{f}_{\ell,m}}) \left(\frac{\partial e_{c,\ell z}}{\partial \mathbf{g}_{\ell,m}} \right)^\dagger = \text{diag}(\overline{\mathbf{f}_{\ell,m}}) \mathbf{v}_{\ell,m}, \quad (5.12)$$

where

$$\begin{aligned}\mathbf{v}_{\ell,M} &= \left(\frac{\partial e_{c,lz}}{\partial \mathbf{g}_{\ell,M}} \right)^\dagger = -\mathbf{H}_{\Delta z_{M,\ell},\lambda_{\text{ex}}}^\dagger \mathbf{H}_{z,\lambda_{\text{ex}}}^\dagger \mathbf{H}_c^\dagger \text{diag} \left(\frac{\mathbf{H}_c \mathbf{H}_{z,\lambda_{\text{ex}}} \mathbf{G}_\ell}{|\mathbf{H}_c \mathbf{H}_{z,\lambda_{\text{ex}}} \mathbf{G}_\ell|} \right) \mathbf{e}_{c,lz} \\ \mathbf{v}_{\ell,m} &= \left(\frac{\partial e_{c,lz}}{\partial \mathbf{g}_{\ell,m+1}} \cdot \frac{\partial \mathbf{g}_{\ell,m+1}}{\partial \mathbf{g}_{\ell,m}} \right)^\dagger = \mathbf{H}_{\Delta z_{m,\ell},\lambda_{\text{ex}}}^\dagger \text{diag}(\overline{\mathbf{t}_{m+1}}) \mathbf{v}_{\ell,m+1}, \quad m = 1, \dots, M-1,\end{aligned}\quad (5.13)$$

are auxiliary vectors for intermediate gradient derivation steps, \dagger is the Hermitian operation, and $\bar{\cdot}$ is the complex conjugate operation. With these auxiliary vectors, it is relatively simple to derive the gradient of the speckle field vector, \mathbf{p}_c , as

$$\nabla_{\mathbf{p}_c} e_{c,lz} = \left(\frac{\partial e_{c,lz}}{\partial \mathbf{g}_{\ell,1}} \cdot \frac{\partial \mathbf{g}_{\ell,1}}{\partial \mathbf{p}_c} \right)^\dagger = \mathbf{S}_\ell^\dagger \mathbf{Q}^\dagger \mathbf{H}_{\Delta s_\ell,\lambda_{\text{ex}}}^\dagger \text{diag}(\overline{\mathbf{t}_1}) \mathbf{v}_{\ell,1}. \quad (5.14)$$

As for the gradient of the pupil function, $\tilde{\mathbf{h}}_c$, we have

$$\nabla_{\tilde{\mathbf{h}}_c} e_{c,lz} = \left(\frac{\partial e_{c,lz}}{\partial \tilde{\mathbf{h}}_c} \right)^\dagger = -\text{diag}(\overline{\mathbf{F}\mathbf{G}_\ell}) \text{diag}(\overline{\tilde{\mathbf{h}}_{z,\lambda_{\text{ex}}}}) \mathbf{F} \text{diag} \left(\frac{\mathbf{H}_c \mathbf{H}_{z,\lambda_{\text{ex}}} \mathbf{G}_\ell}{|\mathbf{H}_c \mathbf{H}_{z,\lambda_{\text{ex}}} \mathbf{G}_\ell|} \right) \mathbf{e}_{c,lz} \quad (5.15)$$

Gradient derivation for 3D fluorescence imaging

To optimize Eq. (5.6) for $\mathbf{o}_1, \dots, \mathbf{o}_M, \mathbf{p}_c, \tilde{\mathbf{h}}_f$, we need to take the derivative of the fluorescent cost function with respect to each. First, we express the gradient for the fluorescence distribution vectors from different layers, $\mathbf{o}_1, \dots, \mathbf{o}_M$ as

$$\nabla_{\mathbf{o}_m} e_{f,\ell} = \left(\frac{\partial e_{f,\ell}}{\partial \mathbf{o}_m} \right)^\dagger = -2 \text{diag}(|\mathbf{f}_{\ell,m}|^2) \mathbf{K}_{z_{m,\ell}}^\dagger \mathbf{e}_{f,\ell}, \quad m = 1, \dots, M \quad (5.16)$$

Then, we would like to derive the gradient of the speckle field, \mathbf{p}_c , as

$$\nabla_{\mathbf{p}_c} e_{f,\ell} = \sum_{m=1}^M \left(\frac{\partial e_{f,\ell}}{\partial \mathbf{f}_{\ell,m}} \cdot \frac{\partial \mathbf{f}_{\ell,m}}{\partial \mathbf{p}_c} \right)^\dagger = -2 \sum_{m=1}^M \left(\frac{\partial \mathbf{f}_{\ell,m}}{\partial \mathbf{p}_c} \right)^\dagger \text{diag}(\mathbf{f}_{\ell,m}) \text{diag}(\mathbf{o}_m) \mathbf{K}_{z_{m,\ell}}^\dagger \mathbf{e}_{f,\ell}, \quad (5.17)$$

where

$$\begin{aligned}\left(\frac{\partial \mathbf{f}_{\ell,m}}{\partial \mathbf{p}_c} \right)^\dagger &= \left(\frac{\partial \mathbf{f}_{\ell,m}}{\partial \mathbf{g}_{\ell,m-1}} \cdot \frac{\partial \mathbf{g}_{\ell,m-1}}{\partial \mathbf{g}_{\ell,m-2}} \dots \frac{\partial \mathbf{g}_{\ell,2}}{\partial \mathbf{g}_{\ell,1}} \frac{\partial \mathbf{g}_{\ell,1}}{\partial \mathbf{p}_c} \right)^\dagger \\ &= \mathbf{S}_\ell^\dagger \mathbf{Q}^\dagger \mathbf{H}_{\Delta s,\lambda_{\text{ex}}}^\dagger \left[\text{diag}(\overline{\mathbf{t}_1}) \mathbf{H}_{\Delta z_1,\lambda_{\text{ex}}}^\dagger \right] \dots \left[\text{diag}(\overline{\mathbf{t}_{m-1}}) \mathbf{H}_{\Delta z_{m-1},\lambda_{\text{ex}}}^\dagger \right].\end{aligned}\quad (5.18)$$

As for the gradient of the pupil function at the fluorescent wavelength, $\tilde{\mathbf{h}}_f$, we can express as

$$\begin{aligned}\nabla_{\tilde{\mathbf{h}}_f} e_{f,\ell} &= -2 \sum_{m=1}^M \text{diag}(\overline{\tilde{\mathbf{h}}_{z_{m,\ell},\lambda_{\text{em}}}}) \mathbf{F} \text{diag}(\mathbf{F}^{-1} \text{diag}(\tilde{\mathbf{h}}_{z_{m,\ell},\lambda_{\text{em}}}) \tilde{\mathbf{h}}_f) \\ &\quad \mathbf{F}^{-1} \text{diag}(\overline{\mathbf{F} \text{diag}(|\mathbf{f}_{\ell,m}|^2) \mathbf{o}_m}) \mathbf{F} \mathbf{e}_{f,\ell}\end{aligned}\quad (5.19)$$

5.2.4 Reconstruction algorithm

5.2.4.1 Initialization of the variables

Since we use a gradient-based algorithm to solve, we must initialize each output variable, ideally as close as possible to the solution, based on prior knowledge.

For 3D coherent reconstructions, the targeted variables are transmittance function, $t_m(\mathbf{r})$, incident speckle field, $p_c(\mathbf{r})$, and pupil function, $\tilde{h}_c(\mathbf{u})$. We have no prior knowledge of the transmittance function or pupil function, so we set $t_m(\mathbf{r}) = 1$ for $m = 1, \dots, M$ and $\tilde{h}_c(\mathbf{u})$ to be a circle function with radius defined by $\text{NA}_{\text{det}}/\lambda_{\text{ex}}$. This initializes with a completely transparent sample and non-aberrated system. If the sample is mostly transparent, the amplitude of our incident speckle field is the overlay of all the in-focus shifted coherent intensities:

$$p_c^{\text{initial}}(\mathbf{r}) = \sum_{\ell=1}^{N_{\text{img}}} I_{c,\ell,z=0}(\mathbf{r} + \mathbf{r}_\ell) / N_{\text{img}}. \quad (5.20)$$

For 3D fluorescence reconstruction, the targeted variables are sample fluorescence distribution, $o_m(\mathbf{r})$, incident field, $p_c(\mathbf{r})$, and pupil function at the emission wavelength, $\tilde{h}_f(\mathbf{u})$. We have no prior knowledge of the system's aberrations, so we set $\tilde{h}_f(\mathbf{u})$ to be a circle function with radius defined by $\text{NA}_{\text{det}}/\lambda_{\text{em}}$. For the incident speckle field, we use the estimated speckle field from the coherent reconstruction as our initialization. The key to a successful 3D fluorescence reconstruction with this dataset is an initialization of the sample's 3D fluorescence distribution using the correlation-based SIM solver [49, 124–127, 130] that gives us an approximate result to start with. We adapt the correlation-based solver in our case for rough 3D SR fluorescence estimation. The basic idea is that we use the knowledge of illumination speckle intensity from the coherent reconstruction to compute the correlation between the speckle intensity and our fluorescence measurement. This correlation is stronger when the speckle intensity lines up with the fluorescent light generated by this excitation in the measurement. Each layer of the estimated speckle intensity gates out out-of-focus fluorescent light in the measurement, so we could get a rough estimate of the 3D fluorescent sample. Mathematically, we express this correlation as

$$o_m^{\text{initial}}(\mathbf{r}) = \sum_{n=1}^9 \left\langle \left(I_{f,\ell}(\mathbf{r}) - \langle I_{f,\ell}(\mathbf{r}) \rangle_{\ell(n)} \right) \left(|f_{m,\ell}(\mathbf{r})|^2 - \langle |f_{m,\ell}(\mathbf{r})|^2 \rangle_{\ell(n)} \right) \right\rangle_{\ell(n)}, \quad (5.21)$$

where $\langle \cdot \rangle_{\ell(n)}$ is the averaging operation over ℓ index of fluorescence images with the same z -scan position (at n -th layer) in the set defined $\ell(n) = \{12^2(n-1) + 1, \dots, 12^2 n\}$.

To understand why this correlation gives a good estimation of the 3D fluorescent sample, we go through a more detailed derivation with a short-hand notation Δ to denote the

operation $\Delta a_\ell(\mathbf{r}) = a_\ell(\mathbf{r}) - \langle a_\ell(\mathbf{r}) \rangle_{\ell(n)}$. Then, we examine one component of Eq. (5.21):

$$\begin{aligned} \langle \Delta |f_{m,\ell}(\mathbf{r})|^2 \Delta I_{f,\ell}(\mathbf{r}) \rangle_{\ell(n)} &= \sum_{m'=1}^M \int o_{m'}(\mathbf{r}') \langle \Delta |f_{m,\ell}(\mathbf{r})|^2 \Delta |f_{m',\ell}(\mathbf{r}')|^2 \rangle_{\ell(n)} h_{f,z_{m'},\ell}(\mathbf{r} - \mathbf{r}') d\mathbf{r}' \\ &\approx \sum_{m'=1}^M \int o_{m'}(\mathbf{r}') \langle \Delta |f_{m,\ell}(\mathbf{r})|^2 \rangle_{\ell(n)}^2 \delta_{m,m'} \delta(\mathbf{r} - \mathbf{r}') h_{f,z_{m'},\ell}(\mathbf{r} - \mathbf{r}') d\mathbf{r}' \propto \langle \Delta |f_{m,\ell}(\mathbf{r})|^2 \rangle_{\ell(n)}^2 o_m(\mathbf{r}), \end{aligned} \quad (5.22)$$

where we assume the speckle intensity is completely uncorrelated spatially in 3D, which is an approximation because the speckle has finite grain size depending on the illumination NA. Under this assumption, this correlation is almost the 3D fluorescence distribution with an extra modulation factor. Hence, it serves well as a initialization for our 3D fluorescence distribution.

5.2.4.2 Iterative algorithm

With all the initializations, we summarize our algorithm for both imaging contrast with the following pseudo-code:

Algorithm 4 3D coherent imaging reconstruction

Require: $I_{c,\ell z}$, \mathbf{r}_ℓ , $\ell = 1, \dots, N_{\text{img}}$

- 1: initialize $t_1^{(1,0)}, \dots, t_M^{(1,0)}, p_c^{(1,0)}, \tilde{h}_c^{(1,0)}$; normalize $I_{c,\ell z}$
 - 2: **for** $k = 1 : K_c$ **do**
 - 3: Sequential gradient descent
 - 4: **for** $j = 1 : (N_{\text{img}} \cdot N_z)$ **do**
 - 5: $z = z_{\text{mod}(j,2)}$; $\ell = \text{mod}(j, N_{\text{img}})$
 - 6: **if** $j < N_{\text{img}} \cdot N_z$ **then**
 - 7: **for** $m = 1 : M$ **do**
 - 8: $t_m^{(k,j)} = t_m^{(k,j-1)} - \nabla_{t_m} e_{c,\ell z}(t_1^{(k,j-1)}, \dots, t_M^{(k,j-1)}, p_c^{(k,j-1)}, \tilde{h}_c^{(k,j-1)}) / 4 \max \left(\left| p_c^{(k,j-1)} \right| \right)^2$
 - 9: **end for**
 - 10: $p_c^{(k,j)} = p_c^{(k,j-1)}(\mathbf{r}) - \nabla_{p_c} e_{c,\ell z}(t_1^{(k,j-1)}, \dots, t_M^{(k,j-1)}, p_c^{(k,j-1)}, \tilde{h}_c^{(k,j-1)}) / \max \left(\left| t_1^{(k,j-1)} \right|, \dots, \left| t_M^{(k,j-1)} \right| \right)^2$
 - 11: $\xi = \mathcal{F}\{G_\ell^{(k,j-1)}\}$
 - 12: $\tilde{h}_c^{(k,j)} = \tilde{h}_c^{(k,j-1)} - \nabla_{\tilde{h}_c} e_{c,\ell z}(t_1^{(k,j-1)}, \dots, t_M^{(k,j-1)}, p_c^{(k,j-1)}, \tilde{h}_c^{(k,j-1)}) \cdot |\xi| / [\max(|\xi|) \cdot (|\xi|^2 + \delta)]$,
 where δ is chosen to be small
 - 13: **else**
 - 14: Do the same update but save to $t_1^{(k+1,0)}, \dots, t_M^{(k+1,0)}, p_c^{(k+1,0)}, \tilde{h}_c^{(k+1,0)}$
 - 15: **end if**
 - 16: **end for**
 - 17: Filter $t_1^{(k+1,0)}, \dots, t_M^{(k+1,0)}$ with Gaussian filter to damp down high-frequency artifacts
 - 18: **end for**
-

Algorithm 5 3D fluorescence imaging reconstruction

Require: $I_{f,\ell}$, \mathbf{r}_ℓ , $f_{1,\ell}, \dots, f_{M,\ell}$ $\ell = 1, \dots, N_{\text{img}}$

- 1: initialize $o_1^{(1,0)}, \dots, o_M^{(1,0)}$, $p_c^{(1,0)}$, $\tilde{h}_f^{(1,0)}$; normalize $I_{f,\ell}$
- 2: **for** $k = 1 : K_f$ **do**
- 3: Sequential gradient descent
- 4: **for** $\ell = 1 : N_{\text{img}}$ **do**
- 5: **if** $j < N_{\text{img}}$ **then**
- 6: **for** $m = 1 : M$ **do**
- 7: $o_m^{(k,j)} = o_m^{(k,j-1)} - \nabla_{o_m} e_{f,\ell}(o_1^{(k,j-1)}, \dots, o_M^{(k,j-1)},$
 $p_c^{(k,j-1)}, \tilde{h}_f^{(k,j-1)}) / \max(|f_{1,\ell}^{(k,j-1)}|, \dots, |f_{M,\ell}^{(k,j-1)}|)^4$
- 8: **end for**
- 9: $p_c^{(k,j)} = p_c^{(k,j-1)}(\mathbf{r}) - \nabla_{p_c} e_{f,\ell}(o_1^{(k,j-1)}, \dots, o_M^{(k,j-1)},$
 $p_c^{(k,j-1)}, \tilde{h}_f^{(k,j-1)}) / \max(|t_1^{(k,j-1)}|, \dots, |t_M^{(k,j-1)}|)^2$
- 10: $\tilde{h}_f^{(k,j)} = \tilde{h}_f^{(k,j-1)} - \nabla_{\tilde{h}_f} e_{f,\ell}(o_1^{(k,j-1)}, \dots, o_M^{(k,j-1)},$
 $p_c^{(k,j-1)}, \tilde{h}_f^{(k,j-1)}) / \max(|\mathcal{F}\{o^{(k,j-1)} \cdot |f_\ell^{(k,j-1)}|^2\}|)^2$
- 11: **else**
- 12: Do the same update but save to $o_1^{(k+1,0)}, \dots, o_M^{(k+1,0)}, p_c^{(k+1,0)}, \tilde{h}_f^{(k+1,0)}$
- 13: **end if**
- 14: **end for**
- 15: **end for**

3D coherent reconstruction takes about 40 iterations, while the 3D fluorescence reconstruction takes around 25 iterations to reach convergence.

5.3 Experimental results

Figure 5.1 shows the experimental setup. A green laser beam (BeamQ, 532 nm, 200 mW) is collimated through a single lens and illuminates the layered Scotch tape element, creating a speckle pattern at the sample. The Scotch tape and the sample are mounted on a 3-axis closed-loop piezo-stage (Thorlabs, MAX311D) and a 1-axis open-loop piezo-stage (Thorlabs, NFL5DP20), respectively, to enable lateral speckle scanning and axial sample scanning. The transmitted diffracted and fluorescent light from the sample then travels through the subsequent $4f$ system formed by the objective lens (Nikon, CFI Achrom 20 \times , NA=0.4) and a tube lens. The coherent and fluorescent light have different wavelengths and are optically separated by a dichroic mirror (Thorlabs, DMLP550R), after which the fluorescence is further spectrally filtered before being imaged onto Sensor-F. The coherent light is ND-filtered and then split by a beam-splitter onto two sensors. Sensor-C1 is in focus, while Sensor-C2 is defocused by 3 mm.

Successful reconstruction relies on appropriate choices for the scanning range and step size [151]. Generally, the translation step size should be 2-3 \times smaller than the targeted resolution and the total translation range should be larger than the diffraction-limited spot size of the original system. Our system has detection NA of 0.4 and targeted resolution of 500 nm, so a 36×36 Cartesian scanning path with a step size of 180 nm is appropriate for 2D SR reconstruction. For coherent imaging, since there is zero axial bandwidth in the coherent TF (Fig. 5.2(a)), the sample's complete diffraction information is projected axially and encoded in the measurement. This enables SR reconstruction of the sample's 3D quantitative phase from just the translating speckle, as in the 2D case. Incoherent imaging, however, demonstrates some degree of optical sectioning due to its torus-shaped TF (Fig. 5.2(b)) - hence fluorescent light that is outside of the DOF of the imaging objective will have weak contrast. Thus, in order to reconstruct 3D fluorescence with high fidelity, we add axial scanning to our acquisition procedure [129].

Here, we use an interleaving scanning scheme, alternating between axial sample scanning and lateral speckle scanning (Fig. 5.1). The 36×36 Cartesian speckle scanning path is divided into 9 blocks of 12×12 sub-scanning paths. Each sub-scanning path is associated with a z -scan position. This means the distance from incident speckle field to sample is

$$\Delta s_\ell = (n - 1)s, \text{ for } \ell = 12^2(n - 1) + 1, \dots, 12^2n, \text{ where } n = 1, \dots, 9, \quad (5.23)$$

where s is the axial step size. We set the fifth z -scan position as the middle of the sample. The total scanning range is roughly the thickness of the sample and the step size is at least $2\times$ smaller than the Nyquist-limited axial resolution of the fluorescence microscope. This extra z -scan enables 3D SR fluorescence information.

5.3.1 3D super-resolution demonstration

With a 0.4 NA objective, our system's native lateral resolution is $1.33 \mu\text{m}$ for coherent imaging and 760 nm for fluorescence (Table 5.1). According to TF analysis, the intrinsic DOF is infinite for coherent imaging and $7.3 \mu\text{m}$ for fluorescence imaging. In order to characterize the resolution capability of our method, we begin by imaging a sample with features below both diffraction limits - a mono-layer of fluorescent polystyrene microspheres with diameter 700 nm . We scan with a z -scan step size of $1 \mu\text{m}$ across $8 \mu\text{m}$ range, fully covering the thickness of the sample. 15 axial layers are assigned to the transmittance function, with each separated by $1.7 \mu\text{m}$ based on Nyquist sampling of the expected axial resolution for our 3D reconstruction.

Figure 5.4 shows that our 3D reconstructions clearly resolve the sub-diffraction individual microspheres and demonstrate better sectioning ability in both coherent and fluorescent channels compared to standard widefield imaging. In the reconstruction, the average lateral peak-to-peak distance of these microspheres is around 670 nm , which is smaller than the nominal size of each microsphere. This is likely due to vertical staggered stacking of the microspheres. Given that our lateral resolution is at least 670 nm , we do break the lateral diffraction limit for both coherent and fluorescent channels, and the coherent channel

achieves $> 2\times$ lateral resolution improvement. Axially, we demonstrate $6\ \mu\text{m}$ resolution for both channels, which is beyond the axial diffraction limit for both channels. The coherent channel improves the axial resolution from no sectioning ability to $6\ \mu\text{m}$, suggesting that the illumination NA of this speckle is >0.4 .

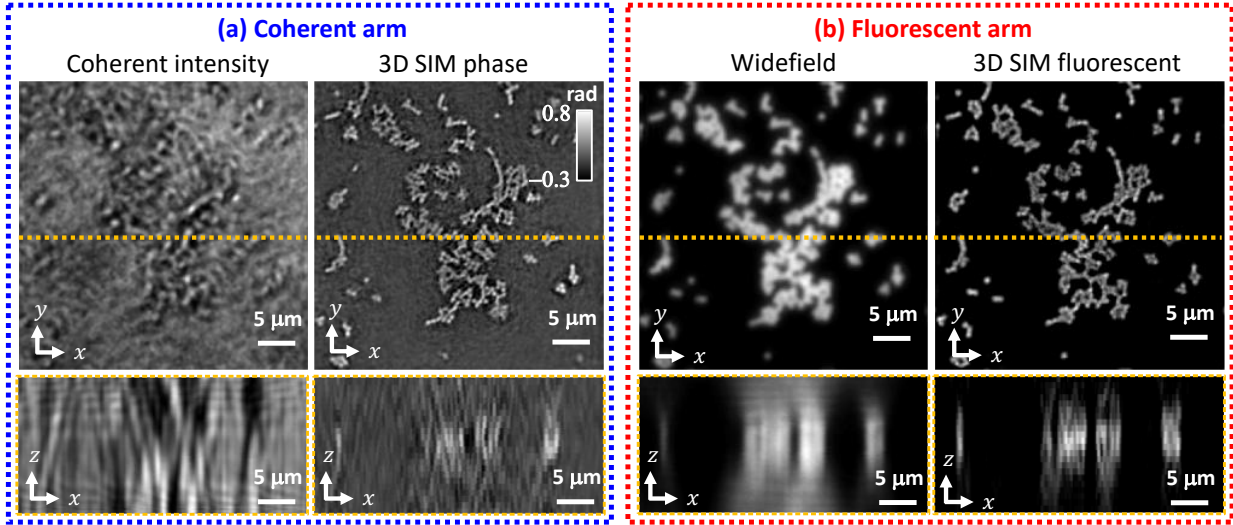


Figure 5.4: 3D multimodal (fluorescence and phase) SIM reconstruction compared to wide-field fluorescence and coherent intensity images for $700\ \text{nm}$ fluorescent microspheres. Resolution beyond the system’s diffraction limit is achieved in both the (a) coherent and (b) fluorescent arms.

5.3.2 3D large-FOV multimodal demonstration

Next we demonstrate 3D multimodal imaging for our full sensor area, giving a FOV $\sim 314\ \mu\text{m} \times 500\ \mu\text{m}$. The native system resolution ($1.33\ \mu\text{m}$ for coherent and $760\ \text{nm}$ for fluorescence) corresponds to ~ 1.4 Mega-pixels of information. As characterized previously, our method achieves $\sim 0.6 \times 0.6 \times 6\ \mu\text{m}^3$ resolution for both fluorescence and phase imaging over an axial range of $\sim 24\ \mu\text{m}$. This corresponds to ~ 14 Mega-voxels of information, a $\sim 10\times$ increase in the information amount. Our experiments are only a prototype; this technique should be scalable to the Gigavoxel range with a higher-throughput objective and higher illumination NA.

Figure 5.5 shows the full-sensor 3D quantitative phase and fluorescence reconstructions of a multi-size sample (mixed $2\ \mu\text{m}$, $4\ \mu\text{m}$ fluorescent, and $3\ \mu\text{m}$ non-fluorescent polystyrene microspheres). We zoom-in on 2 regions of interest (ROIs), displaying 4 axial layers for each. The arrows highlight $2\ \mu\text{m}$ fluorescent microspheres, which defocus more quickly than the larger microspheres. The locations of the fluorescent microspheres match well in both the fluorescence and phase channels. However, there are some locations in the fluorescence

reconstruction where $4\ \mu\text{m}$ microspheres collapse because the immersing media is dissolving the beads over time.

Finally, we demonstrate our technique on human colorectal adenocarcinoma (HT-29) cells fluorescently tagged with AlexaFluor phalloidin, which labels F-actin filaments (sample preparation details in Sec. 5.3.3). During acquisition, we use a z -scan step size of $1.6\ \mu\text{m}$ across a $12.8\ \mu\text{m}$ range. We assign 19 axial layers to the transmittance function, separated by $1.7\ \mu\text{m}$. Figure 5.6 shows the full-sensor 3D quantitative phase and fluorescence reconstructions, with zoom-ins on 2 ROIs. The sample's morphological features, as visualized with quantitative phase, match well with the F-actin visualization of the fluorescent channel. This is expected since F-actin filaments are generally known to encapsulate the cell body.

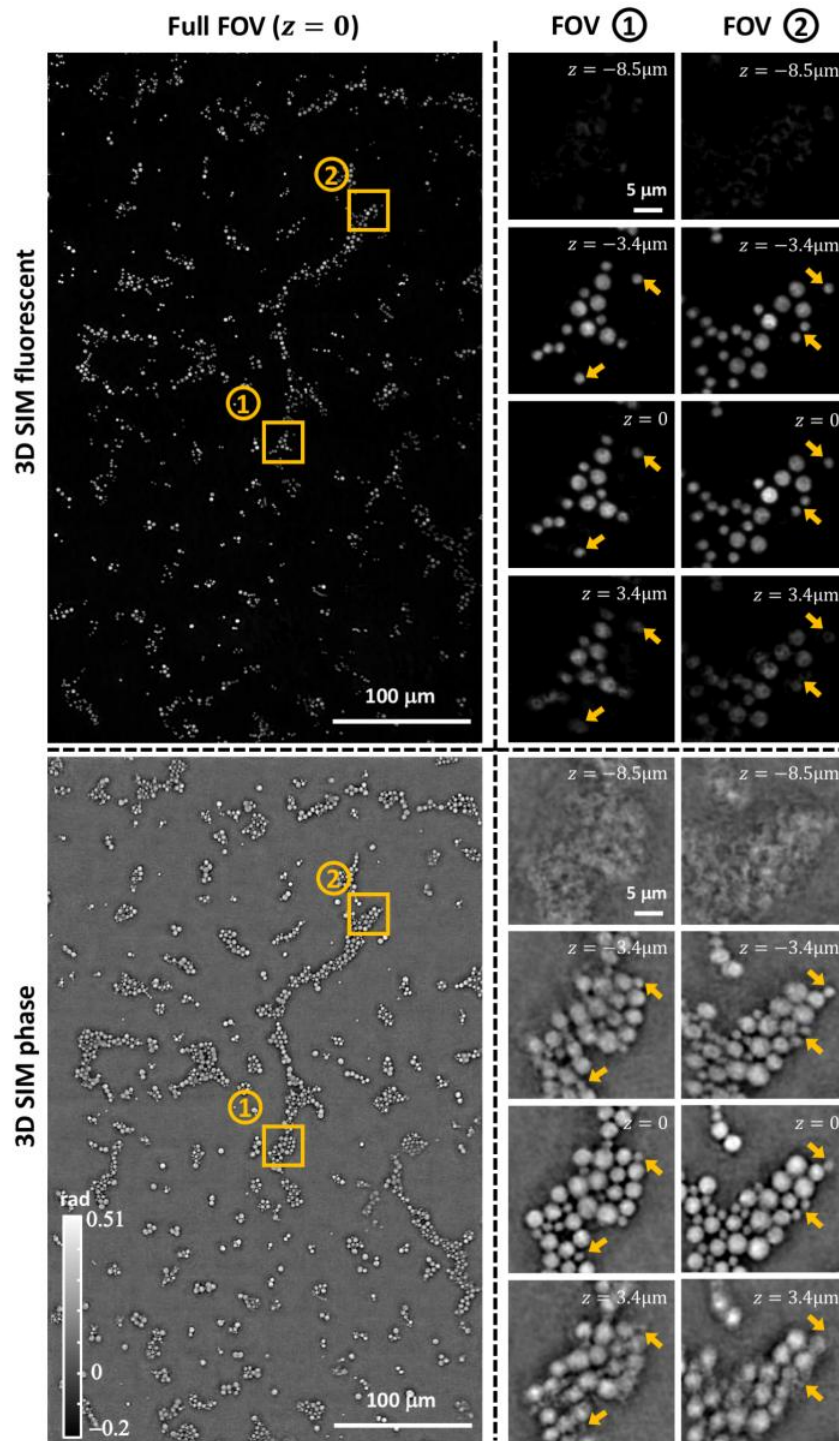


Figure 5.5: Reconstructed 3D multimodal (fluorescence and phase) large-FOV for mixed 2 μm , 4 μm fluorescent and 3 μm non-fluorescent polystyrene microspheres. Zoom-ins for two ROIs show fluorescence and phase at different depths.

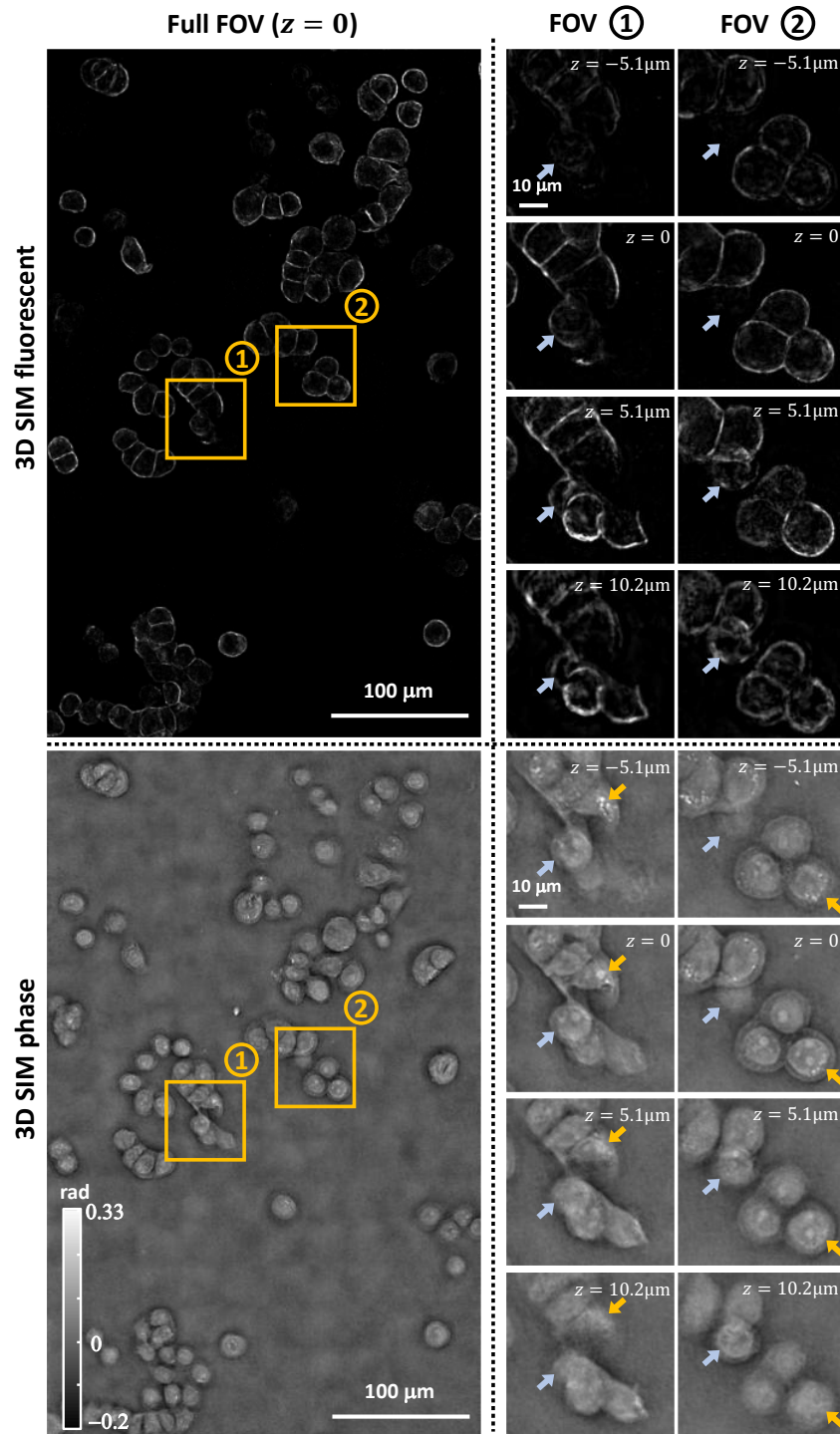


Figure 5.6: Reconstructed 3D multimodal (fluorescence and phase) large-FOV imaging for HT-29 cells. Zoom-ins for two ROIs show fluorescence and phase at different depths. The blue arrows in two ROIs indicate two-layer cell clusters that come in and out of focus. The orange arrows indicate intracellular components, including higher-phase-contrast lipid vesicles at $z = -5.1 \mu\text{m}$, nucleolus at $z = 0$, as well as the cell nucleus and cell-cell membrane contacts.

5.3.3 Sample preparation

The sample shown in Fig. 5.4 is a monolayer of 700 nm diameter polystyrene microspheres (Thermofischer, R700), prepared by placing microsphere dilutions (60 uL stock-solution/500 uL isopropyl alcohol) onto #1.5 coverslips and then allowing to air dry. Water is subsequently placed on the coverslip to reduce the index-mismatch of the microspheres to the air. An adhesive spacer followed by another #1.5 coverslip was placed on top of the original coverslip to assure a uniform sample layer for imaging.

The sample used in Fig. 5.5 is a mixture of 2 μm (Thermofischer, F8826) and 4 μm (Thermofischer, F8858) fluorescently-tagged ($\lambda_{\text{em}} = 605 \text{ nm}$) and 3 μm non-fluorescent (Sigma-Aldrich, LB30) polystyrene microspheres. We follow a similar procedure as before, except that the dilution is composed of 60 uL stock solution of each type of microspheres and 500 uL isopropyl alcohol. Since the microspheres are larger in size, we adopt high-index oil ($n_m(\lambda) = 1.52$ at $\lambda = 532 \text{ nm}$) for sample immersion.

Figure 5.6 uses a sample of HT-29 cells grown in DMEM with 10% FBS, trypsonized with $1\times$ trypsin, passaged twice a week into 100mm dishes at 1/5, 1/6, 1/8 dilutions and stored in a 37C 5% CO_2 incubator. For imaging, HT-29 cells were grown on glass coverslips (12mm diameter, No. 1 thickness; Carolina Biological Supply Co.) and fixed with 3% paraformaldehyde for 20min. Fixed cells were blocked and permeabilized in phosphate buffered saline (PBS; Corning Cellgro) with 5% donkey serum (D9663, Sigma-Aldrich), 0.3% Triton X-100 (Fisher Scientific) for 30 minutes. Cells were incubated with Alexa Fluor 546 Phalloidin (A22283, ThermoFisher Scientific) for 1 hour, washed 3 times with PBS, and mounted onto a second glass coverslip (24x50mm, No. 1.5 thickness; Fisher Scientific) and immobilized with sealant (Cytoseal 60; Thermo Scientific).

5.4 Discussion

In this work, we successfully demonstrate 3D resolution improvement for both quantitative phase and fluorescent imaging. Unlike traditional 3D SIM or 3D quantitative phase methods which use expensive spatial-light-modulators (SLMs) [14, 142] or galvanometer/MEMs mirrors [9, 139, 143] for programmed illumination, our technique is relatively simple and inexpensive. Layered Scotch tape efficiently creates speckle patterns with $\text{NA}_{\text{illum}} > 0.4$, which is hard to achieve with traditional patterning approaches to high-content imaging (*e.g.* lenslet array or grating masks [23–29]). Furthermore, the random structured illumination conveniently multiplexes both phase and fluorescence information into the system’s aperture, enabling us to achieve 3D multimodal SR.

We note that our fluorescent reconstruction relies on the recovered 3D speckle from the coherent imaging channel - mismatch between the two channels can result in artifacts that degrade resolution. The SR gain we achieve in the fluorescent channel does not match that achieved in the coherent channel - we attribute this mainly to mismatch in axial alignment between the coherent and fluorescent imaging cameras. The long DOF of the imaging objec-

tive made it difficult to axially align the cameras to within the axial resolution limit of the high-resolution speckle illumination pattern. In addition, our 3D coherent reconstruction suffers from coherent noise due to system instabilities during the long acquisition process. Specifically, 3D phase information is encoded into the speckle-like (high dynamic range) features within the measurements, which are affected by Poisson noise. These factors reduce performance in both the 3D phase and fluorescence reconstructions.

Another limitation is the relatively long acquisition time - ~ 1200 translations of the Scotch tape results in ~ 180 seconds (without hardware optimization). The number of acquisitions could potentially be reduced with further investigation of the scanning range, step size, and optimal trajectory. This would also reduce computational processing time for the reconstruction, which currently takes ~ 6 hours on a NVIDIA, TITAN Xp GPU with MATLAB, for each $40 \times 40 \mu\text{m}^2$ patch. Cloud computing could in future parallelize the reconstruction by patches.

5.5 Conclusion

We have presented a 3D SIM multimodal (phase and fluorescence) technique using Scotch tape as the patterning element. The Scotch tape efficiently generates high-resolution 3D speckle patterns over a large volume, which multiplexes 3D super-resolution phase and fluorescence information into our low-NA imaging system. A computational optimization algorithm based on 3D coherent and incoherent imaging models is developed to both solve the inverse problem and self-calibrate the unknown 3D random speckle illumination and the system's aberrations. The result is 3D sub-diffraction fluorescence reconstruction and 3D sub-diffraction phase reconstruction with $> 2\times$ lateral resolution enhancement. The method is potentially scalable for Gigavoxel imaging.

Chapter 6

Conclusion and future work

This thesis has explored various coherent and incoherent imaging methods for super-resolution microscopy and high-content microscopy. The fundamental theme in this thesis is to relax the traditional optical design problem for more flexible imaging with the help of optimization algorithms. Optimization formulation serves as a more general framework than analytic inversion for information retrieval, which helps to solve problems with more complex nonlinear information mapping. This approach has already triggered research on unexplored imaging methods and demonstrated several improvements on the existing methods. Throughout this thesis, our focus is to apply this approach for super-resolution coherent and incoherent imaging. The main contributions include:

- Analysis of various phase retrieval optimization algorithms in the context of the super-resolution coherent imaging technique, Fourier ptychographic microscopy (FPM), and discovery of the important role of the noise model in experimental robustness of the reconstruction.
- Development of a self-calibration algorithm for FPM to correct for experimental errors such as system aberrations and light source misalignment.
- Development and experimental verification of a blind algorithm for structured illumination microscopy (SIM) to achieve better and more robust 2D and 3D super-resolution fluorescence imaging with random illumination patterns.
- Proposal and experimental verification of a high-content ($4\times$ super-resolution gain over the intrinsic system) fluorescence and phase microscopy framework with a speckle structured illumination approach. Self-calibration algorithms were developed in tandem to improve the imaging quality.
- Introduction of a general 3D imaging model for the speckle SIM framework. Development and experimental verification of the corresponding algorithm for 3D super-resolution phase and fluorescence imaging with random speckle.

Beyond this work, there are many unexplored directions to use optimization for better solutions of inverse problems or better design of imaging systems. We list a few potential directions that are most related to the current work and worth exploring in the near future.

- **Engineered speckle pattern:** In this work, there are many examples of using random speckle illuminations for coherent and incoherent super-resolution imaging. Theoretically, the multiplexing capability of super-resolution information is highly dependent on the properties of speckle patterns such as the distribution of spatial frequency content and underlying statistics of the speckles. As [78, 160, 161] suggest, it is possible to control properties of the speckle pattern such as the statistics and correlation. A future work combining the idea of engineered speckle patterns and our imaging method would be worth pursuing. For example, to extend our work in Chapter 4 for super high-content imaging, we need to increase the illumination NA even more than what we use now ($\text{NA}_{\text{illum}} \approx 0.35$). Illumination NA not only affects the lateral size but also the axial size of the speckle pattern. The higher the illumination NA, the smaller the lateral and axial size of the speckle grain. Hence, the axial interaction of the speckle and the sample becomes more evident, which requires the 3D model discussed in Chapter 5 and deviates from the purpose of simply doing 2D high-content imaging. In order to stick to 2D high-content imaging, we need to have a speckle pattern with small lateral but large axial grain size. Under this condition, we can borrow the concept from Bessel beams [162, 163] to explore and create this type of Bessel speckle pattern as shown in [164].
- **Joint-channel imaging in scattering media:** In Chapter 5, we leveraged the 3D scattering information from the coherent imaging channel to predict 3D speckle intensity for 3D super-resolution fluorescence reconstruction. This demonstrates a use case of our 3D coherent reconstruction for further prediction or modeling. Let us consider a different geometry, where the scattering layer is in between the fluorescent sample and the detection optics. The fluorescent measurement on the camera is blurry because of the scattering layer. A normal deblurring process would require the knowledge of a point spread function for proper deconvolution. If we could set up a coherent imaging channel to use 3D FPM or our 3D method to estimate the 3D scattering structure, we would have better chance to obtain the knowledge of this space-variant point spread function for later deconvolution. This could be a potential way to tackle imaging problem in scattering media.

In addition to those natural continuations of the work described in this thesis, we now discussed two more general situations in computational imaging to which optimization-based approaches can yield better solutions than analytic inversions.

The first is when the targeted information is highly nonlinear in the imaging model so that there is no obvious analytic relationship between that information and the feasible physical measurement. For example, sample scattered electric field (or phase) is nonlinear to the intensity measurement, thus an optimization formulation is required unless there is

a way to turn the measurement linear to the electric field, such as in holography. Previous research has also suggested that using a computational multislice model can be an effective approach for 3D refractive index retrieval [79, 156, 165]. Multiple scattering slices used to model refractive index of 3D samples interact with each other and are nonlinear with intensity. In this thesis, we also investigate these examples, including FPM, blind SIM (a nonlinear problem because the pattern and the sample are unknown variables), 2D and 3D multimodal SIM with an unknown pattern. Imaging modalities where imaging contrast is a highly nonlinear function of the desired solution are the best candidates for improved solutions using computational imaging. For example, nonlinear label-free microscopy, such as coherent anti-Stokes Raman scattering (CARS) microscopy [166] and stimulated Raman scattering (SRS) microscopy [167], measure intensity that is a nonlinear function of the desired quantity. With a proper computational model, more measurements that contain more information such as resolution or 3D information can potentially give better solutions. A similar example is polarization microscopy [168], where asymmetric refractive index is encoded in the spatial distribution of the 3D electric field that is nonlinear with the intensity measurement. The joint-estimation problem used for self-calibration throughout this thesis is a nonlinear problem as well. Thus, computational imaging approaches using self-calibration can be applied to a variety of imaging systems to achieve better imaging quality.

The second situation is the optimal acquisition design problem. Since a lot of computational imaging problems are nonlinear optimization problems, there is no easy way to analyze how information is transferred in the measurement. Many times people intuitively come up with one or two ways to acquire desired information with limited understanding and reasoning on optimality of the acquisition scheme. There is a strong need to understand how to better acquire data with computational imaging techniques. Up to now, this type of design problem relied on a framework called unrolled optimization [169, 170]. This framework starts with a set of ground truth images and then it generates measurements based on some acquisition schemes. By comparing reconstructions to its corresponding ground truths, the appropriateness of the current acquisition scheme can be evaluated. It will be a future step of the computational imaging community to start thinking about this design problem for every computational imaging technique.

As mentioned, it is difficult to analyze the performance and well-posedness of computational imaging techniques. Usually, thorough simulations and experimental verification are required to justify the effectiveness of reconstruction and characterize the convergence conditions of the technique. For example, thorough simulations and experiments have been done to verify the reconstruction effectiveness of multiplexed FPM [21] as well as our methods in Chapter 4 and 5. In addition, the measurements are usually nonlinear to the targeted unknown variables. This also makes resolution analysis more difficult than methods based on direct analytic linear inversion. Having a systematic way to analyze the well-posedness and performance of computational imaging techniques not only is important challenge in its own right but will also aid in making system design more efficient.

Bibliography

- [1] E. Abbe. “Beitrage zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung”. In: *Arch. Mikroskop Anat.* 9 (1873), pp. 413–420.
- [2] W. Lukosz. “Optical systems with resolving powers exceeding the classical limit. II”. In: *J. Opt. Soc. Am.* 57 (1967), pp. 932–941.
- [3] S. W. Hell and J. Wichmann. “Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy”. In: *Opt. Lett.* 19 (1994), pp. 780–782.
- [4] M. A. A. Neil, R. Juškaitis, and T. Wilson. “Method of obtaining optical sectioning by using structured light in a conventional microscope”. In: *Opt. Lett.* 22 (1997), pp. 1905–1907.
- [5] R. Heintzmann and C. Cremer. “Laterally modulated excitation microscopy: improvement of resolution by using a diffraction grating”. In: *Proc. SPIE* 3568 (1999), pp. 185–196.
- [6] M. G. L. Gustafsson. “Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy”. In: *Journal of Microscopy* 198.2 (2000), pp. 82–87.
- [7] E. Betzig et al. “Imaging intracellular fluorescent proteins at nanometer resolution”. In: *Science* 313 (2006), pp. 1642–1645.
- [8] M. J. Rust, M. Bates, and X. Zhuang. “Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)”. In: *Nature Methods* 3.10 (2006), pp. 793–795.
- [9] Y. Sung et al. “Optical diffraction tomography for high resolution live cell imaging”. In: *Opt. Express* 17.1 (2009), pp. 266–277.
- [10] Y. Choi et al. “Overcoming the diffraction limit using multiple light scattering in a highly disordered medium”. In: *Phys. Rev. Lett.* 107.2 (2011), p. 23902.
- [11] M. Kim et al. “High-speed synthetic aperture microscopy for live cell imaging”. In: *Opt. Lett.* 36 (2011), pp. 148–150.
- [12] Y. Cotte et al. “Marker-free phase nanoscopy”. In: *Nature Photonics* 7 (2013), pp. 113–117.

- [13] S. Chowdhury et al. “Structured illumination multimodal 3D-resolved quantitative phase and fluorescence sub-diffraction microscopy”. In: *Biomed. Opt. Express* 8 (2017), pp. 2496–2518.
- [14] S. Chowdhury et al. “Refractive index tomography with structured illumination”. In: *Optica* 4 (2017), pp. 537–545.
- [15] K. Kim et al. “Correlative three-dimensional fluorescence and refractive index tomography: bridging the gap between molecular specificity and quantitative bioimaging”. In: *Biomed. Opt. Express* 8 (2017), pp. 5688–5697.
- [16] S. Chowdhury et al. “Structured illumination microscopy for dualmodality 3D sub-diffraction resolution fluorescence and refractive-index reconstruction”. In: *Biomed. Opt. Express* 8 (2017), pp. 5776–5793.
- [17] W. Xu et al. “Digital in-line holography for biological applications”. In: *PNAS* 98 (2001), pp. 11301–11305.
- [18] W. Bishara et al. “Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution”. In: *Opt. Express* 18 (2010), pp. 11181–11191.
- [19] A. Greenbaum et al. “Increased space-bandwidth product in pixel super-resolved lens-free on-chip microscopy”. In: *Scientific reports* 3: 1717 (2013).
- [20] G. Zheng, R. Horstmeyer, and C. Yang. “Wide-field, high-resolution Fourier ptychographic microscopy”. In: *Nat. Photon.* 7.9 (2013), pp. 739–745.
- [21] L. Tian et al. “Multiplexed coded illumination for Fourier Ptychography with an LED array microscope”. In: *Biomed. Opt. Express* 5 (2014), pp. 2376–2389.
- [22] L. Tian et al. “Computational illumination for high-speed in vitro Fourier ptychographic microscopy”. In: *Optica* 2.10 (Oct. 2015), pp. 904–911.
- [23] S. Pang et al. “Wide and scalable field-of-view Talbot-grid-based fluorescence microscopy”. In: *Opt. Lett.* 37 (2012), pp. 5018–5020.
- [24] A. Orth and K. Crozier. “Microscopy with microlens arrays: high throughput, high resolution and light-field imaging”. In: *Opt. Express* 20 (2012), pp. 13522–13531.
- [25] A. Orth and K. Crozier. “Gigapixel fluorescence microscopy with a water immersion microlens array”. In: *Opt. Express* 21 (2013), pp. 2361–2368.
- [26] S. Pang et al. “Wide field-of-view Talbot grid-based microscopy for multicolor fluorescence imaging”. In: *Opt. Express* 21 (2013), pp. 14555–14565.
- [27] A. Orth and K. B. Crozier. “High throughput multichannel fluorescence microscopy with microlens arrays”. In: *Opt. Express* 22 (2014), pp. 18101–18112.
- [28] A. Orth et al. “Gigapixel multispectral microscopy”. In: *Optica* 2 (2015), pp. 654–662.
- [29] S. Chowdhury, J. Chen, and J. Izatt. “Structured illumination fluorescence microscopy using Talbot self-imaging effect for high-throughput visualization”. In: *arXiv* 1801.03540 (2018).

- [30] K. Guo et al. “13-fold resolution gain through turbid layer via translated unknown speckle illumination”. In: *Biomed. Opt. Express* 9 (2018), pp. 260–274.
- [31] M. H. Kim et al. “Virtual microscopy as a practical alternative to conventional microscopy in pathology education”. In: *Basic Appl. Pathol.* 1 (2008), pp. 46–48.
- [32] F. R. Dee. “Virtual microscopy in pathology education”. In: *Human Pathol* 40 (2009), pp. 1112–1121.
- [33] R. Pepperkok and J. Ellenberg. “High-throughput fluorescence microscopy for systems biology”. In: *Nat. Rev. Mol. Cell Biol.* 7 (2006), pp. 690–696.
- [34] J. C. Yarrow et al. “Screening for cell migration inhibitors via automated microscopy reveals a Rho-kinase Inhibitor”. In: *Chem. Biol.* 12 (2005), pp. 385–395.
- [35] V. Laketa et al. “High-content microscopy identifies new neurite outgrowth regulators”. In: *Mol. Biol. Cell* 18 (2007), pp. 242–252.
- [36] A. Trounson. “The production and directed differentiation of human embryonic stem cells”. In: *Endocr. Rev.* 27(2) (2006), pp. 208–219.
- [37] U. S. Eggert et al. “Parallel chemical genetic and genome-wide RNAi screens identify cytokinesis inhibitors and targets”. In: *PLoS Biol.* 2 (2004), e379.
- [38] J. W. Goodman. *Introduction to Fourier optics*. Roberts & Co., 2005.
- [39] J. W. Goodman. *Statistical Optics*. Wiley-Interscience, July 2000.
- [40] W. Lukosz. “Optical systems with resolving powers exceeding the classical limit”. In: *J. Opt. Soc. Am.* 56.1463–1472 (1966).
- [41] C. J. Schwarz, Y. Kuznetsova, and S. R. J. Brueck. “Imaging interferometric microscopy”. In: *Opt. Lett.* 28 (2003), pp. 1424–1426.
- [42] V. Mico et al. “Synthetic aperture superresolution with multiple off-axis holograms”. In: *J. Opt. Soc. Am. A* 23.3162–3170 (2006).
- [43] T. Wilson and C. J. R. Sheppard. “Theory and Practice of Scanning Optical Microscopy”. In: *Academic Press* (1984).
- [44] C. J. R. Sheppard. “Scanning Confocal Microscopy”. In: *Encyclopedia of Opt. Engineering* (2003), pp. 2525–2544.
- [45] A. G. York et al. “Resolution doubling in live, multicellular organisms via multifocal structured illumination microscopy”. In: *Nat. Methods* 9 (2012), pp. 749–754.
- [46] A. G. York et al. “Instant super-resolution imaging in live cells and embryos via analog image processing”. In: *Nat. Methods* 10 (2013), pp. 1122–1126.
- [47] F. Ströhl and C. F. Kaminski. “A joint Richardson-Lucy deconvolution algorithm for the reconstruction of multifocal structured illumination microscopy data”. In: *Methods Appl. Fluoresc.* 3 (2015), p. 014002.

- [48] N. Chakrova et al. “Studying different illumination patterns for resolution improvement in fluorescence microscopy”. In: *Opt. Express* 23 (2015), pp. 31367–31383.
- [49] J. García, Z. Zalevsky, and D. Fixler. “Synthetic aperture superresolution by speckle pattern projection”. In: *Opt. Express* 13 (2005), pp. 6075–6078.
- [50] D. Sylman et al. “Random angular coding for superresolved imaging”. In: *Appl. Opt.* 49 (2010), pp. 4874–4882.
- [51] E. Mudry et al. “Structured illumination microscopy using unknown speckle patterns”. In: *Nat. Photon.* 6 (2012), pp. 312–315.
- [52] J. Min et al. “Fluorescent microscopy beyond diffraction limits using speckle illumination and joint support recovery”. In: *Scientific Reports* 3 (2013), 2075:1–6.
- [53] S. Dong et al. “High-resolution fluorescence imaging via pattern-illuminated Fourier ptychography”. In: *Opt. Express* 22.17 (2014), pp. 20856–20870.
- [54] H. Yilmaz et al. “Speckle correlation resolution enhancement of wide-field fluorescence imaging”. In: *Optica* 2.5 (2015), pp. 424–429.
- [55] M. Kim et al. “Superresolution imaging with optical fluctuation using speckle patterns illumination”. In: *Scientific Reports* 5.16525 (2015), p. 16525.
- [56] A. Negash et al. “Improving the axial and lateral resolution of three-dimensional fluorescence microscopy using random speckle illuminations”. In: *J. Opt. Soc. Am. A* 33 (2016), pp. 1089–1094.
- [57] S. Labouesse et al. “Joint reconstruction strategy for structured illumination microscopy with unknown illuminations”. In: *ArXiv: 1607.01980* (2016).
- [58] M. G. L. Gustafsson. “Nonlinear structured-illumination microscopy: wide-field fluorescence imaging with theoretically unlimited resolution”. In: *PNAS* 102 (2005), pp. 13081–13086.
- [59] E. H. Rego et al. “Nonlinear structured-illumination microscopy with a photoswitchable protein reveals cellular structures at 50-nm resolution”. In: *PNAS* 109 (3) (2012), E135–E143.
- [60] D. Li et al. “Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics”. In: *Science* 349 (2015), aab3500.
- [61] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [62] E. Candès, J. Romberg, and T. Tao. “Stable signal recovery from incomplete and inaccurate measurements”. In: *Comm. Pure Appl. Math.* 59 (2006), pp. 1207–1223.
- [63] J. M. Rodenburg and H. M. L. Faulkner. “A phase retrieval algorithm for shifting illumination”. In: *Applied Physics Letters* 85.20 (2004), pp. 4795–4797.
- [64] P. Thibault and M. Guizar-Sicairos. “Maximum-likelihood refinement for coherent diffractive imaging”. In: *New J. Phys.* 14.6 (2012), p. 063004.

- [65] F. Pfeiffer. “X-ray ptychography”. In: *Nat. Photon.* 12 (2018), pp. 9–17.
- [66] Y. Jiang et al. “Electron ptychography of 2D materials to deep sub-angstrom resolution”. In: *Nature* 559 (2018), pp. 343–349.
- [67] H. I. Rasool et al. “Measurement of the intrinsic strength of crystalline and polycrystalline graphene”. In: *Nat. Commun.* 4 (2013), p. 2811.
- [68] D. J. Brady et al. “Compressive Holography”. In: *Opt. Express* 17 (2009), pp. 13040–13049.
- [69] V. Studer et al. “Compressive fluorescence microscopy for biological and hyperspectral imaging”. In: *PNAS* 109 (2012), E1679–E1687.
- [70] G. Zheng, C. Kolner, and C. Yang. “Microscopy refocusing and dark-field imaging by using a simple LED array”. In: *Optics Letters* 36.20 (2011), pp. 3987–3989.
- [71] L. Tian, J. Wang, and L. Waller. “3D differential phase-contrast microscopy with computational illumination using an LED array”. In: *Opt. Lett.* 39.5 (Mar. 2014), pp. 1326–1329.
- [72] D. K. Hamilton and C. J. R. Sheppard. “Differential phase contrast in scanning optical microscopy”. In: *J. Microscopy* 133.1 (1984), pp. 27–39.
- [73] S. B. Mehta and C. J. R. Sheppard. “Quantitative phase-gradient imaging at high resolution with asymmetric illumination-based differential phase contrast”. In: *Opt. Lett.* 34 (2009), pp. 1924–1926.
- [74] T. N. Ford, K. Chu, and J. Mertz. “Phase-gradient microscopy in thick tissue with oblique back-illumination”. In: *Nature Methods* 9.12 (2012), pp. 1195–1197.
- [75] L. Tian and L. Waller. “Quantitative differential phase contrast imaging in an LED array microscope”. In: *Opt. Express* 23 (2015), pp. 11394–11403.
- [76] M. Chen, L. Tian, and L. Waller. “3D differential phase contrast microscopy”. In: *Biomed. Opt. Express* 7 (2016), pp. 3940–3950.
- [77] M. Levoy, Z. Zhang, and I. McDowall. “Recording and controlling the 4D light field in a microscope using microlens arrays”. In: *Journal of Microscopy* 235.2 (2009), pp. 144–162.
- [78] L. Waller, G. Situ, and J. W. Fleischer. “Phase-space measurement and coherence synthesis of optical beams”. In: *Nat. Photon.* 6 (2012), pp. 474–479.
- [79] L. Tian and L. Waller. “3D intensity and phase imaging from light field measurements in an LED array microscope”. In: *Optica* 2 (2015), pp. 104–111.
- [80] G. R. Brady, M. Guizar-Sicairos, and J. R. Fienup. “Optical wavefront measurement using phase retrieval with transverse translation diversity”. In: *Optics express* 17.2 (2009), pp. 624–639.

- [81] A. Williams et al. “Fourier ptychographic microscopy for filtration-based circulating tumor cell enumeration and analysis”. In: *Journal of biomedical optics* 19.6 (2014), p. 066007.
- [82] R. Horstmeyer et al. “Digital pathology with Fourier ptychography”. In: *Computerized Medical Imaging and Graphics* 42 (2015), pp. 38–43.
- [83] J. Chung et al. “Counting White Blood Cells from a Blood Smear Using Fourier Ptychographic Microscopy”. In: *PloS one* 10.7 (2015).
- [84] R. W. Gerchberg and W. O. Saxton. “Phase determination for image and diffraction plane pictures in the electron microscope”. In: *Optik* 34.3 (1971), pp. 275–284.
- [85] J. R. Fienup. “Reconstruction of an object from the modulus of its Fourier transform”. In: *Optics Letters* 3 (1978).
- [86] J. R. Fienup. “Phase retrieval algorithms: a comparison”. In: *Appl. Opt.* 21.15 (1982), pp. 2758–2769.
- [87] V. Elser. “Phase retrieval by iterated projections”. In: *J. Opt. Soc. Am. A* 20.1 (2003), pp. 40–55.
- [88] S. Marchesini. “A unified evaluation of iterative projection algorithms for phase retrieval”. In: *Rev. Sci. Instrum.* 78 (2007), p. 011301.
- [89] M. Guizar-Sicairos and J. M. Fienup. “Phase retrieval with transverse translation diversity: a nonlinear optimization approach”. In: *Optics Express* 16.10 (2008), pp. 7264–7278.
- [90] A. M. Maiden and J. M. Rodenburg. “An improved ptychographical phase retrieval algorithm for diffractive imaging”. In: *Ultramicroscopy* 109 (2009), pp. 1256–1262.
- [91] P. Thibault et al. “Probe retrieval in ptychographic coherent diffractive imaging”. In: *Ultramicroscopy* 109.4 (2009), pp. 338–343.
- [92] C. Yang et al. “Iterative Algorithms for Ptychographic Phase Retrieval”. In: *arXiv* 1105.5628 (2011).
- [93] X. Ou, G. Zheng, and C. Yang. “Embedded pupil function recovery for Fourier ptychographic microscopy”. In: *Opt. Express* 22.5 (2014), pp. 4960–4972.
- [94] E. J. Candès, X. Li, and M. Soltanolkotabi. “Phase Retrieval via Wirtinger Flow: Theory and Algorithms”. In: *arXiv* 1407.1065 (2014).
- [95] L. Bian et al. “Fourier ptychographic reconstruction using Wirtinger flow optimization”. In: *Opt. Express* 23.4 (2015), pp. 4856–4866.
- [96] J. Nocedal and S. Wright. “Numerical Optimization”. In: *Springer Science & Business Media* (Dec. 2006).
- [97] B. Recht, M. Fazel, and P. A. Parrilo. “Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization”. In: *SIAM Review* 52.3 (2010), pp. 471–501.

- [98] E. J. Candès et al. “Phase Retrieval via Matrix Completion”. In: *SIAM Journal on Imaging Sciences* 6.1 (Feb. 2013), pp. 199–225.
- [99] E. J. Candès, T. Strohmer, and V. Voroninski. “PhaseLift: Exact and stable signal recovery from magnitude measurements via convex programming”. In: *Communications on Pure and Applied Math* 66.8 (2013), pp. 1241–1274.
- [100] S. Burer and R. Monteiro. “A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization”. In: *Mathematical Programming* 95 (2003), pp. 329–357.
- [101] R. Horstmeyer et al. “Solving ptychography with a convex relaxation”. In: *New J. Phys.* 17 (2015), p. 053044.
- [102] E. J. Candès, X. Li, and M. Soltanolkotabi. “Phase Retrieval from Coded Diffraction Patterns”. In: *arXiv* 1310.3240 (2013).
- [103] K. Kreutz-Delgado. “The complex gradient operator and the CR-calculus”. In: *arXiv* 0906.4835 (2009).
- [104] J. Nocedal and S. J. Wright. “Conjugate gradient methods”. In: *Springer* (2006).
- [105] A. M. Maiden et al. “An annealing algorithm to correct positioning errors in ptychography”. In: *Ultramicroscopy* 120 (2012), pp. 64–72.
- [106] F. Zhang et al. “Translation position determination in ptychographic coherent diffraction imaging”. In: *Opt. Express* 21.11 (2013), pp. 13592–13606.
- [107] A. Tripathi, I. McNulty, and O. G. Shpyrko. “Ptychographic overlap constraint errors and the limits of their numerical recovery using conjugate gradient descent methods”. In: *Opt. Express* 22.2 (2014), pp. 1452–1466.
- [108] R. Horstmeyer et al. “Overlapped Fourier coding for optical aberration removal”. In: *Optics Express* 22.20 (2014), pp. 24062–24080.
- [109] S. A. Shroff, J. R. Fienup, and D. R. Williams. “Phase-shift estimation in sinusoidally illuminated images for lateral superresolution”. In: *J. Opt. Soc. Am. A* 26 (2009), pp. 413–424.
- [110] S. A. Shroff, J. R. Fienup, and D. R. Williams. “Lateral superresolution using a posteriori phase shift estimation for a moving object: experimental results”. In: *J. Opt. Soc. Am. A* 27 (2010), pp. 1770–1782.
- [111] K. Wicker et al. “Phase optimisation for structured illumination microscopy”. In: *Opt. Express* 21 (2013), pp. 2032–2049.
- [112] K. Wicker. “Non-iterative determination of pattern phase in structured illumination microscopy using auto-correlations in Fourier space”. In: *Opt. Express* 21 (2013), pp. 24692–24701.
- [113] R. Ayuk et al. “Structured illumination fluorescence microscopy with distorted excitations using a filtered blind-SIM algorithm”. In: *Opt. Lett.* 38 (2013), pp. 4723–4726.

- [114] A. Jost et al. “Optical sectioning and high resolution in single-slice structured illumination microscopy by thick slice blind-SIM reconstruction”. In: *PLoS ONE* 10 (2015), e0132174.
- [115] T. Dertinger et al. “Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)”. In: *PNAS* 106.52 (2009), pp. 22287–22292.
- [116] I. J. Cox, C. J. R. Sheppard, and T. Wilson. “Improvement in resolution by nearly confocal microscopy”. In: *Appl. Opt.* 21 (1982), pp. 778–781.
- [117] C. J. R. Sheppard. “Super-resolution in confocal imaging”. In: *Optik* 80 (1988), pp. 53–54.
- [118] C. B. Müller and J. Enderlein. “Image Scanning Microscopy”. In: *Phys. Rev. Lett.* 104 (2010), p. 198101.
- [119] C. J. R. Sheppard, S. B. Mehta, and R. Heintzmann. “Superresolution by image scanning microscopy using pixel reassignment”. In: *Opt. Lett.* 38 (2013), pp. 2889–2992.
- [120] S. Roth et al. “Optical photon reassignment microscopy (OPRA)”. In: *Optical Nanoscopy* 2:5 (2013), pp. 1–6.
- [121] N. Parikh and S. Boyd. “Proximal Algorithms”. In: *Found. Trends Optim.* 1 (2013), pp. 123–231.
- [122] Y. Nesterov. “A method for solving the convex programming problem with convergence rate $O(1/k^2)$ ”. In: *Dokl. Akad. Nauk SSSR* 269 (1983), pp. 543–547.
- [123] A. Beck and M. Teboulle. “A Fast Iterative Shrinkage-Thresholding Algorithm for Linear Inverse Problems”. In: *SIAM J. Imaging Science* 2 (2009), pp. 183–202.
- [124] T. Tanaami et al. “High-speed 1-frame/ms scanning confocal microscope with a microlens and Nipkow disks”. In: *Appl. Opt.* 41 (1996), pp. 4704–4708.
- [125] J. G. Walker. “Non-scanning confocal fluorescence microscopy using speckle illumination”. In: *Opt. Commun.* 189 (2001), pp. 221–226.
- [126] S.-H. Jiang and J. G. Walker. “Experimental confirmation of non-scanning fluorescence confocal microscopy using speckle illumination”. In: *Opt. Commun.* 238 (2004), pp. 1–12.
- [127] R. Heintzmann and P. A. Benedetti. “High-resolution image reconstruction in fluorescence microscopy with patterned excitation”. In: *Appl. Opt.* 45 (2006), pp. 5037–5045.
- [128] L.-H. Yeh and L. Waller. “3D super-resolution optical fluctuation imaging (3D-SOFI) with speckle illumination”. In: *2016 Imaging and Applied Optics Congress* (2016), CW5D.2.
- [129] M. G. L. Gustafsson et al. “Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination”. In: *Biophys J.* 94 (2008), pp. 4957–4970.

- [130] L.-H. Yeh, L. Tian, and L. Waller. “Structured illumination microscopy with unknown patterns and a statistical prior”. In: *Biomed. Opt. Express* 8 (2017), pp. 695–711.
- [131] J. A. Fuerst. “Beyond Prokaryotes and Eukaryotes: Planctomycetes and cell organization”. In: *Nat. Education* 3.9 (2010), p. 44.
- [132] A. W. Lohmann et al. “Space-bandwidth product of optical signals and systems”. In: *J. Opt. Soc. Am. A* 13 (1996), pp. 470–473.
- [133] B. McCullough et al. “Digital microscopy imaging and new Approaches in toxicologic pathology”. In: *Toxicol Pathol.* 32 (2004), pp. 49–58.
- [134] V. Starkuviene and R. Pepperkok. “The potential of high-content high-throughput microscopy in drug discovery”. In: *Br. J. Pharmacol* 152 (2007), pp. 62–71.
- [135] P. Olshausen and A. Rohrbach. “Coherent total internal reflection dark-field microscopy: label-free imaging beyond the diffraction limit”. In: *Opt. Lett.* 38 (2013), pp. 4066–4069.
- [136] S. Chowdhury, A.-H. Dhalla, and J. Izatt. “Structured oblique illumination microscopy for enhanced resolution imaging of non-fluorescent, coherently scattering samples”. In: *Biomed. Opt. Express* 3 (2012), pp. 1841–1854.
- [137] S. Chowdhury and J. A. Izatt. “Structured illumination quantitative phase microscopy for enhanced resolution amplitude and phase imaging”. In: *Biomed. Opt. Express* 4 (2013), pp. 1795–1805.
- [138] P. Gao, G. Pedrini, and W. Osten. “Structured illumination for resolution enhancement and autofocusing in digital holographic microscopy”. In: *Opt. Lett.* 38 (2013), pp. 1328–1330.
- [139] K. Lee et al. “Time-multiplexed structured illumination using a DMD for optical diffraction tomography”. In: *Opt. Lett.* 42 (2017), pp. 999–1002.
- [140] M. Schürmann et al. “Three-dimensional correlative single-cell imaging utilizing fluorescence and refractive index tomography”. In: *J. Biophoton.* (2017), e201700145.
- [141] S. Shin et al. “Super-resolution three-dimensional fluorescence and optical diffraction tomography of live cells using structured illumination generated by a digital micromirror device”. In: *arXiv* 1801.00854 (2018).
- [142] R. Förster et al. “Simple structured illumination microscope setup with high acquisition speed by using a spatial light modulator”. In: *Opt. Express* 22 (2014), pp. 20663–20677.
- [143] D. Dan et al. “DMD-based LED-illumination Super-resolution and optical sectioning microscopy”. In: *Scientific Reports* 3 (2013), p. 1116.
- [144] L.-H. Yeh et al. “Experimental robustness of Fourier ptychography phase retrieval algorithms”. In: *Opt. Express* 23 (2015), pp. 33213–33238.

- [145] L. Bottou. “Large-scale machine learning with stochastic gradient descent”. In: *International Conference on Computational Statistics* (2010), pp. 177–187.
- [146] N. Streibl. “Phase imaging by the transport equation of intensity”. In: *Opt. Commun.* 49.1 (1984), pp. 6–10. ISSN: 0030-4018. DOI: 10.1016/0030-4018(84)90079-8.
- [147] M. Guizar-Sicairos, S. T. Thurman, and J. R. Fienup. “Efficient subpixel image registration algorithms”. In: *Opt. Lett.* 33 (2008), pp. 156–158.
- [148] R. Horstmeyer et al. “Diffraction tomography with Fourier ptychography”. In: *Optica* 3 (2016), pp. 827–835.
- [149] R. Ling et al. “High-throughput intensity diffraction tomography with a computational microscope”. In: *Biomed. Opt. Express* 9.2130–2141 (2018).
- [150] M. Jang et al. “Wavefront shaping with disorder-engineered metasurfaces”. In: *Nat. Photon.* 12 (2018), pp. 84–90.
- [151] L.-H. Yeh, S. Chowdhury, and L. Waller. “Computational structured illumination for high-content fluorescence and phase microscopy”. In: *Biomed. Opt. Express* 10 (2019), pp. 1978–1998.
- [152] Y. Park et al. “Diffraction phase and fluorescence microscopy”. In: *Opt. Express* 14 (2006), pp. 8263–8268.
- [153] E. Wolf. “Three-dimensional structure determination of semi-transparent objects from holographic data”. In: *Optics Communications* 1 (1969), pp. 153–156.
- [154] V. Lauer. “New approach to optical diffraction tomography yielding a vector equation of diffraction tomography and a novel tomographic microscope”. In: *J. Microscopy* 205 (2002), pp. 165–176.
- [155] M. Debailleul et al. “Holographic microscopy and diffractive microtomography of transparent samples”. In: *Meas. Sci. Technol.* 19 (2008), p. 074009.
- [156] J. M. Cowley and A. F. Moodie. “The scattering of electrons by atoms and crystals. I. A new theoretical approach”. In: *Acta Crystallographica* 10.10 (1957), pp. 609–619.
- [157] C. J. R. Sheppard et al. “Three-dimensional transfer functions for high-aperture systems”. In: *J. Opt. Soc. Am. A* 11 (1994), pp. 593–598.
- [158] M. Gu. *Advanced optical imaging theory*. Ed. by 1st ed. Springer, 2000.
- [159] M. Debailleul et al. “High-resolution three-dimensional tomographic diffractive microscopy of transparent inorganic and biological samples”. In: *Opt. Lett.* 34 (2009), pp. 79–81.
- [160] Y. Bromberg and H. Cao. “Generating Non-Rayleigh Speckles with Tailored Intensity Statistics”. In: *Phys. Rev. Lett.* 112 (2014), p. 213904.
- [161] N. Bender et al. “Customizing speckle intensity statistics”. In: *Optica* 5 (2018), pp. 595–600.

- [162] J. Durnin and J. J. Miceli Jr. “Comparison of Bessel and Gaussian beams”. In: *Opt. Lett.* 13 (1988), pp. 79–80.
- [163] T. A. Planchon et al. “Rapid three-dimensional isotropic imaging of living cells using Bessel beam plane illumination”. In: *Nat. Methods* 8 (2011), pp. 417–423.
- [164] D. D. Battista et al. “From amorphous speckle pattern to reconfigurable Bessel beam via wavefront shaping”. In: *arXiv: 1511.04964* (2015).
- [165] A. M. Maiden, M. J. Humphry, and J. M. Rodenburg. “Ptychographic transmission microscopy in three dimensions using a multi-slice approach”. In: *J. Opt. Soc. Am. A* 29 (2012), pp. 1606–1614.
- [166] J.-X. Cheng, A. Volkmer, and X. S. Xie. “Theoretical and experimental characterization of coherent anti-Stokes Raman scattering microscopy”. In: *J. Opt. Soc. Am. B* 19 (2002), pp. 1363–1375.
- [167] C. W. Freudiger et al. “Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy”. In: *Science* 322 (2008), pp. 1857–1861.
- [168] M. Shribak and R. Oldenbourg. “Techniques for fast and sensitive measurements of two-dimensional birefringence distributions”. In: *Appl. Opt.* 42 (2003), pp. 3009–3017.
- [169] S. Diamond et al. “Unrolled optimization with deep priors”. In: *arXiv: 1705.08041* (2018).
- [170] M. R. Kellman et al. “Physics-based learned design: Optimized coded-Illumination for quantitative phase imaging”. In: *IEEE Trans. Computational Imaging* (2019), TCI.2019.2905434.