Spring 1-1-2015

Hybrid Imaging for Extended Depth of Field Microscopy

Ramzi Nicholas Zahreddine
University of Colorado Boulder, zahreddi@colorado.edu

Follow this and additional works at: https://scholar.colorado.edu/ecn_gradetds

Part of the Electromagnetics and Photonics Commons, Hardware Systems Commons, and the Signal Processing Commons

Recommended Citation
https://scholar.colorado.edu/ecn_gradetds/123

This Dissertation is brought to you for free and open access by Electrical, Computer & Energy Engineering at CU Scholar. It has been accepted for inclusion in Electrical, Computer & Energy Engineering Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact cuscholaradmin@colorado.edu.
Hybrid Imaging for Extended Depth of Field Microscopy

by

Ramzi Nicholas Zahreddine

B.S., Rose-Hulman Institute of Technology, 2009

M.S., University of Colorado at Boulder, 2011

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Department of Electrical, Computer, and Energy Engineering

2016
This thesis entitled:
Hybrid Imaging for Extended Depth of Field Microscopy
written by Ramzi Nicholas Zahreddine
has been approved for the Department of Electrical, Computer, and Energy Engineering

Prof. Carol Cogswell

Prof. Kelvin Wagner

Prof. Stephen Becker

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.
An inverse relationship exists in optical systems between the depth of field (DOF) and the minimum resolvable feature size. This trade-off is especially detrimental in high numerical aperture microscopy systems where resolution is pushed to the diffraction limit resulting in a DOF on the order of 500 nm. Many biological structures and processes of interest span over micron scales resulting in significant blurring during imaging. This thesis explores a two step computational imaging technique known as hybrid imaging to create extended DOF (EDF) microscopy systems with minimal sacrifice in resolution. In the first step a mask is inserted at the pupil plane of the microscope to create a focus invariant system over $10 \times$ the traditional DOF, albeit with reduced contrast. In the second step the contrast is restored via deconvolution. Several EDF pupil masks from the literature are quantitatively compared in the context of biological microscopy. From this analysis a new mask is proposed, the incoherently partitioned pupil with binary phase modulation (IPP-BPM), that combines the most advantageous properties from the literature. Total variation regularized deconvolution models are derived for the various noise conditions and detectors commonly used in biological microscopy. State of the art algorithms for efficiently solving the deconvolution problem are analyzed for speed, accuracy, and ease of use. The IPP-BPM mask is compared with the literature and shown to have the highest signal-to-noise ratio and lowest mean square error post-processing. A prototype of the IPP-BPM mask is fabricated using a combination of 3D femtosecond glass etching and standard lithography techniques. The mask is compared against theory and demonstrated in biological imaging applications.
Dedication

To my friends and family who believed in and supported me along the way.
Acknowledgements

I would like to acknowledge:

My advisor, Prof. Carol Cogswell, my committee members, Prof. Rafael Piestun, Prof. Kelvin Wagner, Prof. Bob McLeod, Prof. Stephen Becker, as well as Prof. Mark Winey and Dr. Shannon Hughes, for their support and feedback throughout the course of my graduate career.

My undergraduate optics professors Prof. Rich Lepkowicz, Prof. Sergio Graneri, and Prof. Rob Bunch for starting me on the path.

My group members, Dr. Jiun-Yann Yu, Simeng Chen, and Jian Xing, as well as my colleagues at the University of Colorado, Dr. Chris Leibs, Dr. Tony Barsic, Dr. Zefram Marks, Martha Bodine, John Pfeiffer, Tzu-Min Ou, and Nikki Sanderson, for the countless conversations that helped me flush out the details.

Kevin Zekis and Ben Frevert for keeping me sane.

My family for their love and support.
## Contents

### Chapter

1. **Introduction and Background**
   - 1.1 Depth of Field in Optical Microscopy ............................................. 3
   - 1.2 Systems for Extending the Depth of Field ......................................... 4
   - 1.3 Review of Pupil Function Engineering for EDF ................................. 8
     - 1.3.1 Amplitude Masks ................................................................. 9
     - 1.3.2 Phase Masks ....................................................................... 11
     - 1.3.3 Complex Masks .................................................................. 16
   - 1.4 Road Map ...................................................................................... 17

2. **Vectorial High NA Imaging Theory** .................................................. 18
   - 2.1 Electromagnetic Waves ................................................................. 18
   - 2.2 Point Spread Function .................................................................. 19
   - 2.3 3D Fourier Optics ....................................................................... 22
     - 2.3.1 Optical Transfer Function ....................................................... 23
     - 2.3.2 Defocus Model .................................................................... 24

3. **Incoherently Partitioned Pupils with Binary Phase Modulation for EDF Microscopy** .................................................. 26
   - 3.1 Qualitative Comparison of EDF Pupil Masks in the Literature ............. 26
     - 3.1.1 Cubic Phase Mask ................................................................. 29
     - 3.1.2 Axicon ............................................................................... 31
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1.3</td>
<td>Binary Phase Modulation</td>
<td>33</td>
</tr>
<tr>
<td>3.1.4</td>
<td>Unbalanced OPD Partitioned Pupils</td>
<td>34</td>
</tr>
<tr>
<td>3.2</td>
<td>Incoherently Partitioned Pupil with BPM</td>
<td>40</td>
</tr>
<tr>
<td>3.3</td>
<td>Design Considerations and Optimization of IPP-BPM Masks</td>
<td>42</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Radii of Pupil Partitions</td>
<td>42</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Number and Location of Binary Phase Transitions</td>
<td>43</td>
</tr>
<tr>
<td>3.4</td>
<td>Off-axis Performance of the IPP-BPM Mask</td>
<td>49</td>
</tr>
<tr>
<td>3.5</td>
<td>Comparison Metrics EDF Pupil Masks</td>
<td>51</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Axial Strehl Ratio</td>
<td>51</td>
</tr>
<tr>
<td>3.5.2</td>
<td>Transverse PSF FWHM</td>
<td>52</td>
</tr>
<tr>
<td>3.5.3</td>
<td>PSF Hilbert Space Angle</td>
<td>52</td>
</tr>
<tr>
<td>3.6</td>
<td>Quantitative Analysis of Increasing Pupil Partitions</td>
<td>53</td>
</tr>
<tr>
<td>3.7</td>
<td>Quantitative Comparison of EDF Pupils</td>
<td>56</td>
</tr>
<tr>
<td>3.8</td>
<td>Conclusions</td>
<td>59</td>
</tr>
<tr>
<td>4</td>
<td>Electronic Detectors and Sources of Noise</td>
<td>61</td>
</tr>
<tr>
<td>4.1</td>
<td>Shot Noise</td>
<td>61</td>
</tr>
<tr>
<td>4.2</td>
<td>Electronic Detectors</td>
<td>62</td>
</tr>
<tr>
<td>4.2.1</td>
<td>CCD</td>
<td>63</td>
</tr>
<tr>
<td>4.2.2</td>
<td>sCMOS</td>
<td>63</td>
</tr>
<tr>
<td>4.2.3</td>
<td>EMCCD</td>
<td>65</td>
</tr>
<tr>
<td>4.3</td>
<td>Quantum Efficiency</td>
<td>67</td>
</tr>
<tr>
<td>4.4</td>
<td>Fixed Pattern Noise</td>
<td>67</td>
</tr>
<tr>
<td>4.5</td>
<td>Spurious Charge</td>
<td>69</td>
</tr>
<tr>
<td>4.6</td>
<td>Readout Noise</td>
<td>70</td>
</tr>
<tr>
<td>4.7</td>
<td>Electron Multiplication Register Noise</td>
<td>71</td>
</tr>
<tr>
<td>4.8</td>
<td>Likelihood Functions</td>
<td>72</td>
</tr>
</tbody>
</table>
Tables

Table

3.1 Simulation parameters for EDF pupil mask comparison. .......................... 27
3.2 PSF comparison metrics for IPP EDF phase mask. ............................... 53
3.3 PSF comparison metrics for IPP-BPM EDF phase mask. ...................... 54
3.4 PSF comparison metrics for a variety of EDF phase masks .................. 57

5.1 ISTA for the TV-Poisson model ...................................................... 80
5.2 FISTA for the TV-Poisson model ..................................................... 80
5.3 Split-Bregman algorithm for the TV-Poisson model .......................... 82
5.4 Improvement in PSNR after deconvolution ...................................... 91
5.5 Split-Bregman algorithm for the TV-Poisson-Gamma model ............... 97

6.1 Comparison of the specified to the fabricated values for the IPP mask .... 104
6.2 Comparison of the specified to the fabricated values for the binary phase mask ... 109
Figures

Figure

1.1 Block diagram of modern optical microscope. ............................................. 1
1.2 Diagram of numerical aperture. ................................................................. 4
1.3 Diagram of a wavefront coded EDF imaging system. .................................. 7
1.4 Family of EDF Pupil Functions. ................................................................. 9
1.5 Axial intensity generated from opaque annuli of increasing radius ............... 10
1.6 Delay of the cubic phase mask. ................................................................. 12
1.7 Diagram of an axicon ................................................................................. 14

2.1 Diagram of the relation between the pupil and focal region. ....................... 21

3.1 Diffraction limited PSF and MTF ............................................................... 28
3.2 Cubic phase mask PSF and MTF ............................................................... 30
3.3 Axicon PSF and MTF .............................................................................. 32
3.4 Binary phase modulated pupil mask ......................................................... 34
3.5 Binary phase modulated pupil mask PSF and MTF .................................... 35
3.6 Incoherently partitioned partitioned pupil mask ........................................ 36
3.7 15-zone incoherently partitioned partitioned pupil PSF and MTF ............... 38
3.8 Comparison matrix for EDF pupil mask properties. ................................... 39
3.9 Diagram of an IPP-BPM mask with 4 pupil partitions and 2-zone BPM ....... 41
3.10 DOF as a function of pupil partitions ...................................................... 43
3.11 IPP-BPM axial intensity profiles with varying amount of ripple ........................................ 45
3.12 DOF as a function of pupil partitions with and without BPM ........................................ 46
3.13 IPP-BPM pupil mask PSF and MTF .................................................................................. 47
3.14 IPP-BPM pupil mask PSF and MTF designed via particle swarm optimization. .......... 48
3.15 Off-axis PSFs of IPP-BPM pupil mask ............................................................................. 50
3.16 Comparison of axial partitioned pupil PSFs to a clear aperture PSF ............................. 55
3.17 Comparison of partitioned pupil transverse PSFs to a clear aperture PSF ................. 55
3.18 Comparison of partitioned pupil PSF Hilbert space angles to a clear aperture PSF .... 56
3.19 Axial PSFs through focus for various phase masks and a clear aperture. .................... 58
3.20 Transverse PSFs through focus for various pupil functions .......................................... 58
3.21 Hilbert space angle comparison of EDF PSFs from various phase masks .................. 59

4.1 Diagram of CCD detector electronics. ................................................................................ 64
4.2 Diagram of sCMOS detector electronics. ........................................................................... 64
4.3 Diagram of EMCCD detector electronics. .......................................................................... 66
4.4 Quantum efficiency curves for various camera types ....................................................... 68

5.1 Simulated Bessel beam EDF PSF for minimization algorithm comparisons. ............. 83
5.2 Equal PSNR cameraman images for FISTA and the split-Bregman algorithm .......... 84
5.3 PSNR and error function vs. $\beta$ ..................................................................................... 86
5.4 Convergence comparison for FISTA and the split-Bregman algorithm ......................... 86
5.5 Ground truth EDF phantom .............................................................................................. 88
5.6 EDF phantoms for various pupil masks prior to deconvolution ................................... 90
5.7 EDF phantoms after deconvolution from an object emitting a peak of 1000 photons 93
5.8 EDF phantoms after deconvolution from an object emitting a peak of 2000 photons 94
5.9 EDF phantoms after deconvolution from an object emitting a peak of 5000 photons 95
5.10 EDF phantoms after deconvolution from an object emitting a peak of 10000 photons 96
5.11 Deconvolution of simulated imaging data from an EMCCD camera ......................... 99
6.1 Sketch of a possible realization of an IPP-BPM phase mask. 101
6.2 Sketch of the actual IPP phase mask that was fabricated. 103
6.3 Micrographs of the actual IPP phase mask that was fabricated. 103
6.4 Surface profiles of the IPP pupil mask. 105
6.5 Fabrication process for BPM mask. 106
6.6 Fabricated BPM mask. 108
6.7 AFM profiles taken from a BPM pupil mask. 111
6.8 Experimental setup for EDF fluorescence microscope. 112
6.9 Simulated and experimental clear aperture PSF comparison. 114
6.10 Best fit simulation to the experimental clear aperture PSF. 115
6.11 Comparison of simulated and experimental IPP PSF. 117
6.12 IPP pupil with pseudo-random phase variations to simulate the observed PSF. 118
6.13 Comparison of simulated and experimental IPP-BPM PSF. 120
6.14 EDF images of a Purkinje neuron axon. 122
6.15 EDF images of a BPAE cells. 123
Chapter 1

Introduction and Background

The earliest compound optical microscope was developed circa 1590 [1] and was reported to have 3-10× magnification. Its use as a scientific instrument, however, was limited by the poor quality of optical glass at the time. In the late 1600s Anton van Leeuwenhoek developed improved techniques in lens fabrication, and made many significant contributions to biology with his microscope. However, it was not until the mid 19th century that, through the combined efforts of Carl Zeiss, Otto Schott, and Ernst Abbe, did the quality of optical glass reach a point where the microscope could become a staple of scientific research [1]. Today, it is an essential tool of biological research, medicine, semiconductor development, and numerous other industries.

Virtually all modern optical microscopes used in academic and industrial settings can be broken down into the following four subsystems illustrated in Fig. 1.1: illumination, imaging, detection, and processing & display. The illumination system creates light and delivers it to the sample that is to be imaged. There are various types of illumination sources which are usually specifically tailored to the given imaging task. For example, a bright field microscope uses an

![Figure 1.1: Block diagram of modern optical microscope.](image-url)
incoherent, broad spectrum light to illuminate a sample. Conversely a scanning confocal microscope utilizes a highly coherent, narrowband laser source to excite fluorescence. Likewise the illumination optics can appear in a number of configurations. In transmission systems a condenser lens focuses light onto the sample, and is opposite of the detection optics. In reflection systems, also known as epi-illumination, the objective lens is used for both illuminating and imaging the sample. The properties of the illumination system can have a dramatic impact on the uniformity, brightness, and resolution of the final image.

The optical system, which traditionally consists of a series of lenses, is responsible for collecting light from the sample, and re-imaging it onto a detector. The configuration of this system can vary greatly from one type of microscope to another, but an infinity corrected objective is usually the primary imaging element in a modern microscope. Light exiting an infinity objective is collimated which allows for additional optics (e.g. beamsplitters, filters, polarizers, prisms, etc.) to be easily introduced into the microscope without affecting the location of the focal plane. The light is then imaged onto the detector by a tube lens. The imaging objective determines the base magnification and resolution of the microscope. These lens systems are precision engineered and fabricated to provide the sharpest image possible given physical constraints.

The function of the modern detector is to convert the optical signal into an electrical signal. Various detectors are available for optical microscopy, the most common of which is the charge coupled device (CCD) camera. In low light conditions scientific complementary metal oxide semiconductor (sCMOS) or electron multiplying CCD (EMCCD) cameras are often used. Advancements in the speed and sensitivity of modern cameras are opening the door to investigate biological processes which were previously unobservable. These devices and their characteristics are discussed in detail in Chapter 4.

The processing and display subsystem takes the electrical signal from the detector, applies any necessary image processing operations, and displays it in the appropriate manner. Recent advancements in image processing and computing power have allowed microscopes to be used in ways previously not possible. Tools such as deconvolution, feature extraction, denoising, contrast
enhancement, and 3D rendering allow the end user to extract and enhance information relevant to his or her specific problem.

The advent of modern cameras, signal processing algorithms, and display technology has spurred the field of computational imaging. The next advancements in computational microscopy will require the purposeful design and integration of all four of these subsystems. The work presented in this thesis utilizes this methodology to overcome the trade-off between depth of field (DOF) and resolution seen in traditional microscopy. Specifically it looks to optimize the performance and capability of extended depth of field (EDF) biological imaging systems.

1.1 Depth of Field in Optical Microscopy

Traditional optical lens design is ever striving to improve resolution. The Rayleigh resolution criterion of an optical system, $\Delta \ell$, is defined as the minimum distance at which the far-field diffraction pattern from two point sources can be definitively distinguished [2]. It is expressed in terms of a system’s numerical aperture (NA) and the wavelength of the light, $\lambda$.

$$\Delta \ell = \frac{0.61 \lambda}{NA}$$

The NA is a measure of the light collecting ability of an optical system. It can be thought of as the largest angle of incident light that can be collected from an in-focus object, as shown in Fig. 1.2. The equation for the NA can be easily derived from the figure and is given as,

$$NA = n_0 \sin(\alpha),$$

where $\alpha$ is the angle of the limiting ray, and $n_0$ is the refractive index of the media between the sample and the optical system. The DOF for a high NA imaging objective was derived in [3, 4] as,

$$DOF = \frac{1.77 \lambda}{4n_0 \sin^2 \frac{\alpha}{2}}.$$  

As can be clearly seen from Eq. 1.3 the NA and DOF of a traditional imaging system have an inverse relationship, which implies that the DOF and the resolution also have an inverse relationship. For
an oil immersion objective with NA = 1.4 and \( n_0 = 1.518 \), imaging green light at the center of the visible spectrum \( \lambda = 550 \text{ nm} \), the DOF is calculated to be 523 nm. Thus, any feature more than 260 nm from the focal plane in either direction will appear significantly blurred. Many important sub-cellular biological processes occur over distances of microns, and therefore in a traditional microscopy system the researcher must choose between a low resolution image that is entirely in focus, or a high resolution image with a limited DOF.

1.2 Systems for Extending the Depth of Field

Several methods and systems have been proposed to overcome the DOF trade-off in optical microscopy. Most systems utilize some combination of scanning, optical design, and image processing to create a reconstructed image that maintains resolution over an extended depth. Several such systems along with their advantages and disadvantages are described below. They can be grouped into the following categories: scanning microscopes, plenoptic microscopes, and wavefront coded (WFC) microscopes. This is by no means an exhaustive list of all available systems, but rather represents a combination of established microscopy techniques along with state of the art methods.

Scanning Microscopy Scanning microscopes are perhaps the most straightforward systems for solving the depth of field problem while maintaining spatial resolution. Confocal scanning microscopy was first proposed by Minsky in 1957 [5]. Laser scanning confocal microscopy (LSCM)
was pioneered by Brakenhoff, Sheppard, Wilson, and a host of others in the late 1970s [6, 7, 8, 9, 10]. It still represents one of the most widely used microscope configurations in the biological sciences today. In it a pinhole is placed in front of the detector (usually a photomultiplier tube) which serves to reject out of focus light. The image is then built up by raster scanning the focused illumination spot over the full 3D region of interest. While this technique produces beautiful images with high a signal-to-noise ratio (SNR), its temporal resolution is limited by the scan time. It also can result in significant photo-bleaching and damage of a sample over the course of a full image acquisition [10]. To improve the acquisition speed spinning disk confocal microscopy was introduced [11, 12], where an array of pinholes rapidly scans the sample plane simultaneously. However, the system suffers from cross talk and very low throughput. Regardless of the exact configuration used, the equipment involved in constructing confocal scanning microscopes is generally quite costly.

More recently selective plane illumination microscopy (SPIM) and related techniques [13, 14, 15] were introduced to optically section samples too thick for traditional confocal microscopy. In this approach a beam is focused into a narrow sheet of light that excites fluorescence in a thin axial section and is then scanned through a sample to acquire 3D information. SPIM performs significantly better than LSCM when imaging thick samples with large refractive index variation. Disadvantages are that it requires extensive sample preparation, expensive opto-mechanical equipment and, while faster than LSCM, is still temporally limited by scanning. Despite these limitations recent advancements along this front have been promising [16].

Another scanning solution to the EDF problem is the wide-field deconvolution microscope [17, 18]. In this approach a full field, through focus scan is acquired, and a 3D reconstruction is created by computationally deconvolving the image. The benefits are that a full field is acquired at each scan point so the temporal resolution is higher than LSCM. However it still suffers from photo-bleaching from multiple exposures and the reconstruction can be quite time consuming as well as computationally difficult. Wide field deconvolution microscopy also generally exhibits lower resolution than confocal systems.

**Light Field Microscopy** A more recent development is that of light field imaging. First
implemented by Ng et al. [19], the light field, or plenoptic, camera utilizes a microlens array placed in front of the detector to record both spatial and angular information. This allows for an image to be digitally refocused over a large DOF. The technique was incorporated into a microscope by Levoy et al. [20, 21] and Lim et al. [22, 23]. The primary disadvantage of a light field imaging system is that spatial resolution is significantly sacrificed in order to obtain the angular resolution. Digital deconvolution and other processing techniques [24, 25, 26] have been proposed to improve the spatial resolution, however they still exhibit inconsistencies through focus and artifacts from the microlens array. Recently, it was also demonstrated that more uniform through focus resolution can be achieved by combining a plenoptic microscope and wavefront coding (described in detail below) [27]. Many of the optical and computational concepts developed in this thesis under the context of wavefront coding can be applied to enhance the quality of plenoptic imaging systems. The two techniques have the potential to benefit greatly from each other in the coming years.

Wavefront Coding The final technique discussed for EDF imaging is known as hybrid imaging or wavefront coding [28, 29], and is the major focus of this thesis. Hybrid imaging as described by Dowski and Cathey is a two step process: The first step is to modulate the pupil so as to encode the optical wavefront with a specific set of information, and the second step is to perform digital image processing to retrieve said information. In the case of EDF imaging the coded information is a blurred, but focus invariant point spread function (PSF) over an elongated focal region. The image processing takes the form of a digital deconvolution to remove the blur and restore contrast to the full EDF region. A diagram of the wavefront coding process for EDF imaging can be seen in Fig. 1.3. Häusler [30] proposed a similar system which utilized purely optical processing techniques as sufficient computational power was not readily available or developed at the time. Beyond the concepts of focus invariance and the use of digital post processing, Dowski and Cathey emphasized two important properties of the optical transfer function (OTF) for WFC systems: it must be invariant through focus and monotonically decrease towards the diffraction limited cutoff frequency. Mino and Okano [31] also noted the importance of these properties in EDF imaging systems.
Figure 1.3: Diagram of a wavefront coded EDF imaging system. In a traditional optical system features beyond the shallow DOF appear significantly blurred (a). The same system with an EDF pupil mask exhibits a constant blurring function through depth (b). A digital deconvolution step is then applied to restore contrast creating an EDF image.
The choice of pupil modulation is the major free variable in the design of WFC systems. Regardless of pupil function, hybrid imaging systems have common properties relevant to their application in optical microscopy. Benefits are that they require no moving parts or through focus scan, and thus have the potential for high speed 3D imaging. Modulation of the pupil can be achieved with elements that can be easily fabricated at relatively low cost [32]. WFC systems can be modularly integrated into existing microscopes. Pupil functions can also be tailored to emphasize the most relevant spatial frequencies for a given application. From a computational perspective WFC systems are much easier to implement than wide field deconvolution microscopy since a single deconvolution filter can be used to reconstruct the EDF image as the PSF is focus invariant. Conversely, a fundamental drawback is that as the DOF region is extended there is an accompanying decrease in the peak signal intensity. This is a product of conservation of energy and its effects on the SNR of the system are described by Silveira et al. [33, 34]. This problem is exacerbated by pupil functions that absorb light. Another inherent drawback of EDF via hybrid imaging is the loss of depth information due to all features in the sample volume appearing sharply focused in the resulting two-dimensional image. Despite these limitations, hybrid imaging for EDF microscopy represents an attractive and simple solution to many experiments where cost, temporal resolution, photo-toxicity, and/or photo-bleaching are primary concerns.

1.3 Review of Pupil Function Engineering for EDF

As mentioned above, the choice of pupil plane modulation is the most important factor in the design of a hybrid imaging system for EDF. In practice this modulation is achieved by placing a mask at, or conjugate to, the pupil plane. To properly guide the design of new and improved pupil masks it is essential to understand the current state of the art. The review that follows has been divided into three subsections: amplitude modulated masks, phase modulated masks, and complex (amplitude and phase) masks as shown in Fig. 1.4. Within each of these subsections there are various pupil function designs, each with its own distinct advantages and disadvantages. It is first necessary to briefly define the pupil function and its role in an imaging system. A complete
theoretical treatment in the context of optical microscopy is presented in Chapter 2. The imaging properties of an optical system can be described by its pupil, \( P(u, v) \).

\[
P(u, v) = A(u, v)e^{-i\phi(u, v)}
\]  

This function describes the amplitude, \( A(u, v) \), and phase, \( \phi(u, v) \), of a given system’s aperture stop, where \( (u, v) \) are the normalized transverse pupil coordinates. From the pupil function one can calculate the PSF and the OTF of the system using the methods of Fourier optics \([2, 35]\). The goal of an EDF hybrid imaging system is to modify the pupil in such a way as to create a PSF that is non-diffracting over a greater axial extent (i.e. focus invariant) and simultaneously create an OTF that features no zeros prior to some cutoff frequency (i.e. no contrast reversals).

### 1.3.1 Amplitude Masks

The effect of pupil plane amplitude modulation on the 3D light field distribution has been studied as far back as the early- to mid-1800s. The trivial implementation of this technique is to reduce the pupil diameter which leads to the resolution trade-off mentioned previously. Airy \([36]\) calculated the on-axis intensity distribution for a lens obscured by a coaxially located opaque disk (i.e. an annular pupil). Lord Rayleigh \([37]\) noted the reduced effect of spherical aberration in the same system and commented on its applications in astronomy. Steward \([38]\) was the first to calculate the effects of an annular aperture on the DOF. Figure 1.5 shows his plots of axial intensity elongating as the size of the opaque disk is increased. Hopkins furthered the subject by calculating the radial
intensity profiles of various annular shaded apertures through focus [39]. Linfoot and Wolf [40] did the same for binary annular apertures.

The works referenced above were largely interested in describing the through focus intensity distribution and characteristics for reflection telescopes where the aperture is obstructed by a secondary mirror. Welford [41] was the first to purposefully describe the use of a binary annular pupil to increase the DOF of a non-astronomical imaging system, specifically those observing sparse, point-like objects. O’Neill [42] calculated the transfer functions for an annular aperture in an aberration free system. Barakat and Houston [43] detailed similar transfer function calculations, but for pupils with an arbitrary amount of spherical aberration. They showed the balancing effect of spherical aberration on the frequency response of the amplitude modulated pupil.

Mino and Okano [31] investigated the DOF effects of shaded annular apodizers with a spherical transverse profile in a traditional photographic imaging system. They showed an improvement in the defocused OTF in comparison to a clear aperture both theoretically and experimentally. Indebetouw and Bai [44] similarly demonstrated EDF imaging with a higher throughput by placing a Fresnel zone plate at the pupil plane. Ojeda-Castañeda, Berriel-Valdos, and Montes characterized the imaging and DOF properties of a Bessel function based on annular apodization.
profiles [45].

In the 1980s the focus shifted from characterization to optimization and design of EDF pupil functions. EDF absorbing apertures were designed using heuristic constraints to minimize the PSF’s variation with defocus [46]. Similarly the Strehl ratio was used as an alternate metric for deriving an absorbing EDF pupil function [47]. Simulations showed that both designs exhibited a greater tolerance to defocus, but offered little in terms of uniformity in the spatial frequency response. Ojeda-Castañeda, Andrés, and Díaz applied similar PSF heuristics to design pupils insensitive to both defocus and spherical aberration [48].

All of the amplitude modulated EDF pupil masks described suffer from two main drawbacks: decreased light throughput and loss of transverse resolution. They are also ill-suited for hybrid imaging systems as the PSF, while elongated, is not particularly focus invariant. This would lead to numerous reconstruction artifacts in the post-processed EDF images. A recently proposed EDF system has attempted to overcome these drawbacks by using a spatial light modulator (SLM) to variably modulate the pupil [49, 50]. However, this technique requires multiple acquisitions to synthesize a full resolution EDF image negating one of the primary benefits of hybrid imaging.

1.3.2 Phase Masks

The 1990s saw a dramatic shift away from amplitude modulated pupil functions for EDF to purely phase modulated functions — the obvious benefit being the increased light throughput. Additionally, these phase mask designs were combined with advancements in computing power and image processing algorithms to achieve some noteworthy results. The phase only modulated EDF systems to be described are divided into two sub categories: radially asymmetric and radially symmetric masks.

Radially Asymmetric Phase Masks Using mathematical tools from radar engineering Dowski and Cathey [29] developed the hybrid imaging methodology for EDF optical systems as described in the previous section. The ambiguity function [51, 52, 53], which maps spatial frequency
information through focus, served as the primary design equation for their phase mask. Assuming a rectangularly separable, nonlinear monomial function for the profile, the stationary phase method [54, 55] was used to derive the cubic phase mask (CPM), whose phase profile is shown in Fig. 1.6. Dowski and Cathey’s analysis was limited to scalar paraxial imaging systems; Arnison calculated the high NA vectorial PSF for the CPM and demonstrated it in a fluorescence microscope [35].

Figure 1.6: Delay of the cubic phase mask.

Other radially asymmetric phase profiles have been designed using comparable base principles and methodologies. The logarithmic phase mask [56] was derived similarly to the CPM, but with fewer restrictions on the functional form of the mask. It slightly outperforms the CPM, but is largely equivalent. The generalized cubic phase mask (GCPM) was developed by allowing higher order nonlinear terms into the phase function and applying a Fisher information metric [57, 58]. The GCPM in general shows a poorer modulation transfer function (MTF), but a better phase transfer function (PTF) than the CPM. A mask with an exponential phase profile was also designed using the Fisher information metric, and again displayed a similar phase profile and performance to the CPM [59].

Asymmetric phase modulated EDF pupil functions offer a great improvement over the purely absorbing filters in that they maintain resolution out to the diffraction limited cutoff fre-
requency, and are much less sensitive to defocus. This enables the digital post-processing to generate EDF images that maintain a comparable resolution to a traditional diffraction limited image. Despite these benefits the asymmetric phase masks lead to some critical negative consequences. The OTF exhibits poorer contrast for spatial frequencies off the u and v axes. Additionally, the PTF of these systems tends to increase with defocus. Thus if the peak-valley phase delay of the mask is not sufficiently large, the PSF shifts away from its true position in the transverse plane with defocus. This was described as a “banana” shaped PSF by Arnison [35]. The accumulation of phase also leads to significant artifacts and quantitative inaccuracies in the post-processed image [60].

**Radially Symmetric Phase Masks** The origin of nearly all radially symmetric EDF phase masks can be traced back to the axicon originally proposed by McLeod [61, 62]. It has a conical shape and, assuming it completely fills a normalized pupil, can be entirely described by its index of refraction and the side angle, δ, as shown in Fig. 1.7. Its EDF properties are easily intuited if the element is thought of as a lens with a variable focal length. The axicon has the effect of smearing light over an axial extent in the form of a quasi-Bessel beam [63]. Bessel beams exhibit various properties that make them appealing for use in EDF microscopy and will be discussed in detail later in this chapter.

The EDF properties were noted by McLeod, but its application in imaging systems went largely undeveloped for the next 25 years. Once research began to focus on axicon type pupils for EDF, it followed a similar path to that of the amplitude apodization and non-rotationally symmetric phase functions. Davidson et al. modified the linear axicon design using a geometrical optics approach to create the axilens [64]. Sochacki et al. argued Davidson’s approach violated the conservation of energy and re-derived the shape in the form of the logarithmic axicon [65, 66]. Using the Wigner distribution [67] (Fourier transform pair of the ambiguity function previously described), Zalvidea and Sicre [68] determined that a quartic phase profile was optimal in a radially symmetric geometry. Their quartic phase mask produced similar results to that of the axilens and logarithmic axicon. Zhou et al. [69] proposed a mask based on radial polynomials named the rational phase mask. It resulted in fewer intensity oscillations than the quartic mask, but still showed a decrease
in spatial frequency content. All four phase masks create quasi-Bessel beams that exhibit minimal
diffraction over an extended DOF. All of the work on radially symmetric pupil masks cited to
this point utilized only the on-axis intensity as their primary design tool and provided no spatial
frequency analysis.

Chi and George [70] were the first to apply radially symmetric pupil masks to digital
imaging systems. They derived a logarithmic axicon from Fermat’s principle, and made use of
hybrid imaging to improve the final resolution. They experimentally measured the cutoff frequency
with defocus and showed a significant increase in DOF, but also a significant decrease in the
maximum achievable resolution despite deconvolution. Zhai et al. demonstrated an axicon in a
hybrid imaging system, although the post processed images were of relatively low quality and again
no formal spatial frequency analysis was performed [71]. Rajesh et al. [72, 73] simulated the EDF
properties of an axicon at the pupil of a high NA imaging objective using vectorial diffraction
theory. Zhao et al. proposed the use of Zemax to design a rational phase mask specifically for
infinity corrected microscope objectives [74]. Cathey and Dowski designed an EDF element similar
to the rational phase mask using Zemax in a similar method to that of Zhao. Known as the
circular caustic phase mask, though never formally published, its profile can be found in [75]. A
number of authors have reported on the use of axicons to structure the illumination and detection

Figure 1.7: Diagram of an axicon with apex angle, δ. The extended depth of field Bessel beam
region is shaded in gray.
in fluorescence imaging with the particular application of scanning two-photon microscopy systems [76, 77].

The axicon based phase masks all utilized a continuous phase profile to extend the DOF, an alternate approach is that of binary diffractive optical elements. These elements consist of annular regions with phase transitions of 0 to \( \pi \), and are referred to as binary phase modulated (BPM) pupil masks herein. Sheppard demonstrated a paraxial design approach for these masks to create maximally flat EDF PSFs [78]. Jabbour et al. demonstrated enhanced optical sectioning using BPM pupil masks under vectorial, high NA assumptions [79, 80]. Liu et al. proposed an annular binary phase mask [81, 82] based off the work of Wang and Gan [83, 84] for simultaneous super resolution and EDF in optical coherence microscopy and endoscopy. Mo proposed combining the quartic and binary phase masks to achieve consistent through focus performance, but at the expense of resolution [85]. While these masks exhibit a large amount of control over the design of the PSF they inherently induce numerous contrast reversals and zeros in the frequency domain, limiting their application in hybrid imaging systems. Their most successful applications have been in structuring the illumination of laser scanning optical systems.

All of the previously described phase masks have been based on manipulating the geometric properties of the pupil. An alternate technique leverages chromatic properties of the light to extend the DOF. This approach uses a tiered annular phase mask to impart an optical path difference (OPD) longer than the coherence length of the light being imaged which effectively partitions the pupil. The resulting PSF mimics that of opaque annular pupils without the loss in throughput. Chu et al. proposed both planar and apsheric versions of these pupil partitioning phase masks for EDF [86, 87, 88]. They were able to extend the DOF of a 60mm focal length system 5× with this method. Abrahamsson et al. proposed a similar system and demonstrated it in fluorescence microscopy [89]. The specific properties of these masks will be discussed in detail in Chapter 3.

Many radially symmetric phase masks offer beneficial properties in the form of “non-diffracting” Bessel beams. The beams self reconstruct and feature a central lobe that is actually narrower than the Airy disk, leading to super resolution in some applications. A major drawback in
many of the axicon based phase masks is that the through focus intensity of the PSF varies greatly over the EDF region. This makes these shapes less useful for doing quantitative analysis in biological imaging. Recent work in tiered annular phase mask designs have improved the uniformity of the axial intensity, but further spatial frequency analysis needs to be done to accurately characterize these pupil functions.

1.3.3 Complex Masks

Modulating both the amplitude and phase of the pupil function offers the most control over the 3D light field distribution. A common method to achieve EDF beams utilizes BPM pupil masks combined with amplitude modulation. Recent research [90, 91] has achieved promising results with this technique and it offers the advantage of relatively straightforward fabrication. Ben-Eliezer et al. demonstrated a number of shaded aperture binary phase masks in widefield electronic imaging systems [92, 93, 94, 95, 96]. However their designs suffer significant resolution loss limiting their applications.

Many groups have utilized complex phase functions in an effort to generate quasi-Bessel beams. Complex holograms for the creation of these PSFs were proposed in [97, 98, 99]. Honkanen and Turunen simulated a two-element system to improve axial intensity uniformity of such beams [100]. Čižmár et al. [101, 102] demonstrated a high degree of intensity control and uniformity by designing holograms via the Gerchberg-Saxton algorithm [103]. The major drawback of the holographic approach is the diffraction efficiency. Some of Čižmár’s implementations achieve only 1% throughput which is not feasible for fluorescence microscopy where photon counts may already be a limiting factor. Due to these limitations, complex pupil masks, similarly to amplitude pupil masks, are not actively considered in this thesis. However, if these systems could be implemented with high efficiency they would have great potential in EDF microscopy.

Despite the extensive amount of literature on the subject of EDF imaging systems, relatively few efforts have been made to incorporate this technique into optical microscopy systems [35, 89]. The main limitations up to this point have been loss of spatial frequency information,
inconsistencies in the through focus PSF, and noise amplification in the post processed image. These limitations are addressed in this thesis by the design of a new type of pupil mask and utilization of recently developed regularized deconvolution algorithms along with state of the art camera technology.

1.4 Road Map

The road map for this thesis is as follows. In Chapter 2 the mathematical framework for image formation in optical microscopy is presented. In Chapter 3 prior art EDF pupil functions are compared in the context of microscopy, and a new EDF pupil function specifically tailored for quantitative biological imaging is described. In Chapter 4 modern camera technology is reviewed and accompanying noise models are presented. In Chapter 5 total variation (TV) regularized deconvolution models are presented for the various imaging conditions along with efficient algorithms for solving said models. In Chapter 6 the newly designed EDF pupil function is fabricated and applications in biological imaging are presented. Conclusions and future work are presented in Chapter 7.
Vectorial High NA Imaging Theory

Vectorial diffraction theory in high NA imaging systems is a well developed subject. The model developed by Richards and Wolf [104] under the Debye approximation [105] provides an efficient method to calculate the spatial electric field distribution near the plane of best focus. The frequency domain interpretation for an arbitrary projected pupil under high NA vectorial assumptions was formalized by Sheppard and Larkin [106], and outlined in the context of EDF microscopy by Arnison et al. [35]. The full model for an arbitrary pupil function is detailed below, for both the spatial and frequency domains. Note there is no assumption of cylindrical symmetry in this formulation which is essential for analyzing asymmetric pupil profiles such as the cubic phase mask.

2.1 Electromagnetic Waves

From Maxwell’s equations in the absence of charge in an isotropic homogeneous medium, one can show that the electric field vector $\mathbf{E}$ must satisfy the steady state Helmholtz wave equation [3],

$$\left( \nabla^2 + k^2 \right) \mathbf{E} = 0,$$

(2.1)

where $k = nk_0 = 2\pi n/\lambda_0$, assuming quasi-monochromatic light, and is commonly referred to as the wave number. Assuming Cartesian components the electric field can be expressed as the following vector $\mathbf{E} = (E_x, E_y, E_z)$. Solutions to the wave equation can be written as an arbitrary electric
field with vectorial amplitude \( E_0 \) and phase \( \phi \) is written as,

\[
E(r) = E_0(r) e^{i \phi(r)}.
\] (2.2)

Both amplitude and phase terms are a function of the Cartesian space vector \( r = (x, y, z) \). A special case of Eq. 2.2 which will be utilized heavily in the imaging model is the plane wave. Assuming unit amplitude it can be written as,

\[
E_p(r) = e^{i k \rho \cdot r},
\] (2.3)

where \( \rho = (u, v, s) \) is a Cartesian pupil space coordinate.

Due to their limited response time, detectors are only sensitive to the intensity of the electric field as the phase varies far too quickly to be observed. The intensity, a scalar quantity, is defined as the square modulus of the electric field vector, \( I = |E|^2 \). In the case of Cartesian components,

\[
I = |E_x|^2 + |E_y|^2 + |E_z|^2.
\] (2.4)

### 2.2 Point Spread Function

The amplitude point spread function (PSF) of an imaging system is defined as 3D distribution of electric field at the focus generated by a plane wave. The electric field near the focus of a high NA imaging system can be expressed as the superposition of plane waves that are modulated by a complex pupil as proposed by Debye in the scalar case [105]. As discussed by Sheppard and Arnison, the Debye approximation is suitable for many cases in high NA microscopy. Richards and Wolf extended Debye’s theory to the vectorial case to yield [104],

\[
E(r) = -\frac{i k}{2\pi} \int \int \int Q(\rho) e^{i k \rho \cdot r} d\rho,
\] (2.5)

where \( Q(\rho) \) is vectorial pupil function of the imaging system as described by McCutchen [107].

For the case of a homogeneous focal region of constant refractive index, the pupil only exists on the surface of a sphere, commonly referred to as Ewald’s Sphere, with a radius defined by the wave number \( k \),

\[
Q(\rho) = Q_s(\rho) \delta (|\rho| - k).
\] (2.6)
This constraint on the pupil function is dictated by Eq. 2.1. Looking only at the forward propagating wave the 3D vectorial pupil function is written as,

\[ Q(\rho) = Q_s(\rho) \delta \left( s - \sqrt{k^2 - w^2} \right), \quad (2.7) \]

where \( w = \sqrt{u^2 + v^2} \).

The axial projected pupil function, \( P'_+(u,v) \), is now introduced as,

\[
P'_+(u,v) = \int_0^\infty Q(\rho) \, ds = Q_s(\rho) \delta \left( s - \sqrt{k^2 - w^2} \right),
\]

\[
= Q_s(u,v,s_+) \frac{1}{s_+}. \quad (2.8)
\]

The projected pupil allows one to represent the 3D pupil as a 2D function by integrating over the axial dimension. As will be shown in the following section, use of the projected pupil greatly simplifies calculation of the PSF. Assuming \( k \) has been normalized to 1, the \( 1/s_+ \) factor accounts for the change in projected thickness of the spherical shell with angle and is defined as \( s_+ = \sqrt{1 - w^2} \).

To account for polarization the electric and magnetic strength vectors \( \mathbf{a}(u,v) \) and \( \mathbf{b}(u,v) \) are respectively introduced as,

\[
\mathbf{a}(u,v) = f E_0 \cos^{1/2} \theta \left[ (\mathbf{g}_0 \cdot \mathbf{i}) \mathbf{g}_1 + (\mathbf{g}_0 \cdot \mathbf{j})(\mathbf{g}_1 \times \rho) \right] \quad (2.9)
\]

\[
\mathbf{b}(u,v) = \left( \frac{\epsilon}{\mu} \right)^{1/2} \mathbf{a}, \quad (2.10)
\]

where \( \epsilon \) is the permittivity of the medium, \( \mu \) is the permeability of the medium, \( f \) is the focal radius of the Gaussian sphere, \( E_0 \) is the amplitude factor, \( \mathbf{g}_0 \) is the unit vector lying in the meridional plane perpendicular to the ray in object space, and \( \mathbf{g}_1 \) is the unit vector lying in the meridional plane perpendicular to the ray in image space. The geometry of the pupil and image space mapping is outlined in Fig. 2.1. The unit vector in the direction of the coordinate axes is denoted by the usual \((\mathbf{i}, \mathbf{j}, \mathbf{k})\). Following the derivation of Richards and Wolf, \( \mathbf{g}_0 \) and \( \mathbf{g}_1 \) can be evaluated as,

\[
\mathbf{g}_0 = \frac{(\rho \times \mathbf{k}) \times \mathbf{k}}{|\rho \times \mathbf{k} \times \mathbf{k}|}, \quad (2.11)
\]
Figure 2.1: Diagram of the relation between the pupil and focal region.

\[ g_1 = \frac{(\rho \times k) \times \rho}{|\rho \times k \times \rho|}, \]  \hfill (2.12)

respectively. Evaluation of Eqs. 2.11 and 2.12 yields the following Cartesian form for the electric and magnetic strength vectors.

\[ \mathbf{a}(u, v) = f E_0 \sqrt{s_+} \begin{pmatrix} (u^2 s_+ + v^2) / w^2 \\ -uv (1 - s_+) / w^2 \\ -u \end{pmatrix} \]  \hfill (2.13)

\[ \mathbf{b}(u, v) = f E_0 \sqrt{s_+} \begin{pmatrix} -uv (1 - s_+) / w^2 \\ (u^2 + v^2 s_+) / w^2 \\ -v \end{pmatrix} \]  \hfill (2.14)

Microscope objectives are traditionally designed to abide by the sine condition in order to achieve aplanatic imaging [108] which is carried in the factor of \( \sqrt{s_+} \) in the above equations. The strength vectors denote polarization of the unperturbed electric and magnetic fields at the pupil. Note that they are \( 90^\circ \) rotations of one another. The corresponding projected pupil functions associated with \( x \) and \( y \) linearly polarized light are.

\[ P'_{+H}(u, v) = \frac{1}{s_+} \mathbf{a}(u, v) T(u, v), \]  \hfill (2.15)
and,
\[ P'_{+V}(u,v) = \frac{1}{s_+} b(u,v) T(u,v), \] (2.16)

where H and V denote horizontal and vertical polarization respectively. The scalar portion of the pupil function is denoted by the transmission function \( T(u,v) \). It contains the complex modulation information for the pupil. For a traditional circularly symmetric optical system in normalized coordinates, \( T(u,v) = \text{circ}(u,v) \).

Revisiting Eq. 2.5 the amplitude PSF near the focus of a high NA imaging objective for a given polarization can now be written as,
\[ E_{(H,V)}(r) = -\frac{ik}{2\pi} \int\int_\Sigma P'_{(H,V)}(u,v) e^{ik\rho r} dudv. \] (2.17)

Note the + subscript will be dropped from all future equations as it is assumed the system only captures the forward propagating wave. Assuming a circularly symmetric pupil the integration area, \( \Sigma \), is limited by the NA of the system and thus extends from \( 0 < w < \sin(\alpha) \). As shown in Eq. 2.4 the intensity PSF is the sum of the square modulus of the individual field components.

The electric field is assumed to be unpolarized unless specified otherwise. The unpolarized intensity PSF is found by averaging the horizontal and vertical linearly polarized intensity PSFs. All of the designs to be presented in thesis are in the context of incoherent imaging and thus linear in intensity, as opposed to coherent imaging systems which are linear in field. Henceforth the intensity PSF shall simply be referred to as the PSF unless explicitly mentioned otherwise.

### 2.3 3D Fourier Optics

Fourier optics provides an ideal framework to analyze both the intensity distribution and spatial frequency effects of an imaging system. Two-dimensional Fourier optics, as outlined by Goodman [2], relates the pupil function, PSF, and OTF through Fourier transforms. However, it relies on the scalar paraxial approximation which is not suitable for the imaging objectives used in high NA microscopy. Through the work of Sheppard, Larkin, and Arnison [35, 106] 2D scalar Fourier optics was adapted to 3D vectorial high NA imaging theory.
2.3.1 Optical Transfer Function

The frequency domain representation of the PSF is the OTF. The 3D OTF describes how spatial frequencies are modified by the system in both amplitude and phase through focus. The OTF is critical to the performance of an EDF system that utilizes deconvolution to improve resolution. In 2D Fourier optics the OTF at the plane of best focus is calculated as the Fourier transform of the PSF at that same plane [2]. The adaptation to 3D begins with the projection slice theorem [109].

\[
f(x, y, 0) \iff \int F(u, v, s) \, ds
\]  

(2.18)

Where the \(\iff\) symbol denotes a Fourier transform pair. The theorem states that a projection in Fourier domain is transformed to a slice in spatial domain. Applying the projected pupil function obtained from Eqs. (2.8), (2.15), and (2.16) to the projection slice theorem yields the amplitude PSF at the plane of best focus.

\[
E_{(H,V)}(x, y, 0) \iff P'_{(H,V)}(u, v)
\]  

(2.19)

The quantity actually measured by the detector however is the intensity PSF. Assuming randomly polarized light, the vectorial intensity PSF at the focal plane, \(h(x, y, 0)\), can be written as the average of the horizontally and vertically polarized intensity PSFs [104],

\[
h(x, y, 0) = \frac{1}{2} \left( I_H(x, y, 0) + I_V(x, y, 0) \right),
\]  

(2.20)

where \(I_H\) and \(I_V\) are defined via Eq. 2.4 as,

\[
I_H(x, y, 0) = |E_{x,H}(x, y, 0)|^2 + |E_{y,H}(x, y, 0)|^2 + |E_{z,H}(x, y, 0)|^2,
\]  

(2.21)

\[
I_V(x, y, 0) = |E_{x,V}(x, y, 0)|^2 + |E_{y,V}(x, y, 0)|^2 + |E_{z,V}(x, y, 0)|^2.
\]  

(2.22)

Up to this point quasi-monochromatic light was assumed. This is obviously an idealization in any experiment, but especially so in fluorescence microscopy. Fluorophores often have a spectral full width half maximum (FWHM), \(\Delta \lambda\), of 40 nm or greater. Thus in order to properly model the PSF across its full spectrum, it is necessary to integrate over all of the wavelengths (or temporal
frequencies) present within the signal. The typical fluorescence emission filter has a bandwidth of approximately 50 nm so this serves to truncate the spectrum regardless of its true width.

\[
h(x, y; 0) = \int S(\nu) \left( \frac{\nu^2}{c^2} \right)^2 \frac{1}{2} \left( I_H(x, y; 0) + I_V(x, y; 0) \right) d\nu
\]  

(2.23)

Assuming an incoherent imaging system the OTF at the focal plane is found by taking the 2D Fourier transform of the intensity PSF.

\[
h(x, y, 0) \rightarrow OTF(u, v, 0)
\]  

(2.24)

The OTF is a complex function that can be written in polar notation. When formulated in this way the magnitude of the OTF is known as the modulation transfer function (MTF), while the argument of its imaginary component is known as the phase transfer function (PTF).

\[
OTF(u, v) = MTF(u, v) e^{iPTF(u, v)}
\]  

(2.25)

For a given optical system, its MTF describes the contrast at which a given spatial frequency is preserved at the image plane. The MTF is a critical to the effectiveness of EDF imaging systems that utilize digital processing. Any zeros in the MTF correspond to a loss of that particular spatial frequency. If a zero should occur before the cutoff frequency a contrast reversal is seen in the regions where the OTF is negative (i.e. a \(\pi\) phase shift in the PTF). An EDF OTF containing zeros and contrast reversals will yield an inaccurate reconstruction and information loss after post-processing.

### 2.3.2 Defocus Model

The 3D Fourier optics adaptation has yielded a method to calculate the PSF and OTF at the plane of best focus from a projection of the 3D pupil function. EDF imaging systems, however, are intended to operate over a region many times the length of a traditional DOF, not just at the focal plane. To investigate the through focus performance of these systems it is necessary to incorporate a defocus term into the model. This can be done easily by representing defocus as an aberration and including it into the projected pupil function. Thus for a defocus of, \(z_d\), the projected pupil
function for an x polarized beam is written as,

$$P'_x(u,v,z_d) = \frac{1}{s_+} a(u,v) T(u,v) T_d(u,v,z_d),$$  \hspace{1cm} (2.26)

where,

$$T_d(u,v,z_d) = \exp(iks_+z_d).$$  \hspace{1cm} (2.27)

The defocused PSF can now be calculated as was done in the previous section. For a translation of $z_d$ along the optical axis, the defocused PSF can be written as,

$$E(x,y,z_d) \Leftrightarrow P'(u,v,z_d).$$  \hspace{1cm} (2.28)

Similarly the defocused OTF can be written in terms of the defocused intensity PSF as,

$$h(x,y,z_d) \Leftrightarrow OTF(u,v,z_d).$$  \hspace{1cm} (2.29)

The advantage of this interpretation is that the fast Fourier transform (FFT) algorithm can be utilized to efficiently calculate the PSF and OTF at a given plane transverse to the optical axis from the projected pupil function.
Chapter 3

Incoherently Partitioned Pupils with Binary Phase Modulation for EDF Microscopy

The main goal of this work is the analysis and optimization of wavefront coded EDF microscopy systems. This chapter serves two main functions towards that goal: it proposes a new family of pupil masks, and establishes quantitative comparison metrics to predict the performance of EDF microscopy systems. As outlined in Chapter 1, the literature is inundated with pupil mask designs that claim EDF properties in one form or another. In order to determine the most effective designs for fluorescence microscopy, several pupil mask designs are taken from the literature and qualitatively evaluated against one another in simulation. Based on this evaluation a new pupil mask is proposed called the incoherently partitioned pupil with binary pupil modulation (IPP-BPM). Some basic properties of the IPP-BPM masks are explored, and initial designs proposed. Finally, the proposed designs are quantitatively compared with the literature using three performance metrics.

3.1 Qualitative Comparison of EDF Pupil Masks in the Literature

Only a handful of the most relevant EDF phase functions from the literature have been included in this comparison. Any pupil masks which feature amplitude modulation were automatically disregarded since most fluorescence applications are already operating in a weak signal regime and cannot afford to throw away photons. Additionally, designs which have been the subject of widespread study were favored. The purpose of these comparisons is to observe general properties of pupil functions to guide in the design of an improved EDF phase mask.
In the following section the PSF, axial intensity, and MTF for each phase mask are simulated and plotted through focus for comparison. All simulations in this section are done numerically in Matlab using the theory outlined in Chapter 2. The parameters used are specified in Table 3.1. The number of transverse samples is denoted by $N_{x,y}$ and the number of axial samples is denoted by $N_z$. The transverse sampling was set to $N_{x,y} = 1024$ as values greater than this significantly increased computation time with no appreciable increase in accuracy. The central wavelength is denoted as $\lambda_c$ and the spectrum is approximated as a Gaussian function with full width half maximum (FWHM) of $\Delta \lambda$.

Table 3.1: Simulation parameters for EDF pupil mask comparison.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_c$</td>
<td>0.508</td>
<td>µm</td>
</tr>
<tr>
<td>$\Delta \lambda$</td>
<td>0.040</td>
<td>µm</td>
</tr>
<tr>
<td>$z_d$</td>
<td>±10</td>
<td>µm</td>
</tr>
<tr>
<td>$N_{x,y}$</td>
<td>1024</td>
<td></td>
</tr>
<tr>
<td>$N_z$</td>
<td>301</td>
<td></td>
</tr>
<tr>
<td>NA</td>
<td>1.40</td>
<td></td>
</tr>
<tr>
<td>$n_0$</td>
<td>1.518</td>
<td></td>
</tr>
</tbody>
</table>

The PSF, axial intensity distribution, and MTF from a clear aperture are shown in Fig. 3.1 and serve as a reference point for all of the EDF pupil functions to follow. The FWHM of the axial PSF is 478 µm which coincides with the 483 µm DOF predicted by Eq. 1.3. The axial intensity highlights some smaller secondary intensity peaks, but the vast majority of the energy is localized within the predicted DOF. The MTF correspondingly shows full spatial frequency content over the same region. Note that the quantity being plotted in Fig. 3.1 is actually $\log_{10}(\text{MTF})$ to aid in visualization. This will be the case for all following MTF plots shown unless noted otherwise. As expected, the MTF decays quickly with defocus and many zeros (black lines) can be seen. Spatial frequencies at the zeros are entirely lost, while spatial frequencies beyond a given zero are contrast reversed until the next zero is reached.
Figure 3.1: (a) Diffraction limited x-z PSF and axial intensity, (b) diffraction limited MTF along $k_x$. 
3.1.1 Cubic Phase Mask

The cubic phase mask (CPM) has been widely studied over the past 20 years [29]. The phase profile can be seen in Fig. 1.6, and its transmission function at the pupil plane can be written as,

\[ T(u, v) = \exp\left(ikA(u^3 + v^3)\right). \]  

The PSF and “axial” intensity, for the cubic are shown in Fig. 3.2(a) for a value of \(A = 100\) which results in a peak-valley phase delay of 25.8 waves. It can be clearly seen that the cubic PSF persists with limited diffraction over an axial range of approximately 7 \(\mu m\), approximately 15 times that of a standard DOF. Unfortunately the PSF has been splayed out along the x (and y) axis, significantly decreasing the peak intensity through focus. Note that the cubic PSF translates through focus due to a linear accumulation of phase, because of this effect the axial intensity plot is actually the maximum projected intensity for this case in order to make a fair comparison with other phase masks. The highly asymmetric PSF means that the intermediate (or encoded) image must be deconvolved to obtain recognizable images. The MTF plot, Fig. 3.2(b), shows that the CPM maintains high frequency content over a wide range of defocus along the \(k_x\) (and \(k_y\)) axis. There are no zeros and it maintains relatively good contrast. The magnitude of the MTF off of the \(k_x\) and \(k_y\) axes is significantly lower however. It still does not feature any zeros before the cutoff frequency, but falls into the noise floor much more rapidly. Since biological objects are rarely rectilinear this could pose problems when observing most specimens.

While this design has some distinct advantages, the asymmetric PSF and frequency response, along with the translation with defocus, limit the effectiveness of the CPM in quantitative imaging experiments such as particle tracking. The accumulation of phase with defocus is also responsible for inherent post-processing artifacts upon deconvolution [60]. Despite these disadvantages it is important to note the desirable properties, and study how they can be incorporated into more symmetric designs. In addition to the MTF possessing no zeros, the other major benefit is that the PSF is axially symmetric about the traditional plane of best focus. This means the pupil mask can be inserted and removed from the system without needing to refocus, which is a
very useful property when setting up experiments. Finally the PSF shows a fairly constant peak intensity through focus which ensures a uniform photometric response regardless of an emitter’s location within the extended DOF.

Figure 3.2: (a) Cubic x-z PSF and maximum projected intensity, (b) Cubic MTF along $k_x$. 
3.1.2 Axicon

The axicon, first proposed by McLeod, is a conical optical element with a linear phase ramp, as shown in Fig. 1.7 [61, 62]. It is defined by its side angle $\delta$ with a transmittance function of

$$T(w) = \exp \left( ikA w \right).$$

Axicons, and their variants [64, 65, 66, 68, 69], create a “non-diffracting” Bessel beam. The PSF for an axicon with $A = 140$ is shown in Fig. 3.3(a). For a pupil 7.2 mm in diameter and assuming an index of refraction of 1.46, this corresponds to an angle $\delta = 0.378^\circ$. The resulting peak-valley phase delay is approximately 21.5 waves. The axicon does have optical power so the PSF is now axially asymmetric about the clear aperture focal plane and shifted several microns in the positive direction. The simulation region has been adjusted accordingly in order to account for this shift. The axial intensity of the PSF grows rapidly, then slowly decays over a range of 16+ $\mu$m. The central lobe of the quasi-Bessel beam broadens with defocus. This coupled with the decaying intensity limits the usable range closer to 7 $\mu$m. The axial intensity also oscillates with a low modulation depth over this range. The MTF, Fig. 3.3(b), shows virtually full frequency content at the beginning of the EDF region. The cutoff frequency then begins to decay with the decreasing intensity and broadening central lobe.

The axicon features a few benefits over the CPM, the first being that the intensity is more concentrated along the optical axis. Another benefit is that the rotationally symmetric PSF and MTF result in intermediate and processed images which have a uniform spatial frequency response regardless of feature orientation. A drawback of the axicon is the slowly decaying PSF beyond the usable DOF region. This results in fluorescing objects outside the extended DOF adding to the background intensity and thus reduces contrast. Additionally, the decaying MTF results in significant resolution loss towards the edge of the DOF. Lastly, the shift in the PSF from the traditional plane of best focus makes their use cumbersome when actually performing experiments.
Figure 3.3: (a) Axicon x-z PSF and axial intensity intensity, (b) Axicon MTF along $k_x$. 
Binary Phase Modulation

More recently, binary phase modulation (BPM) has been a subject of numerous studies and demonstrated to create non-diffracting beams of uniform intensity. Figure 3.4 shows a diagram of a pupil plane with BPM. The annular modulation alternates between values of 0 (blue regions) and $\pi$ (gray regions). In the paraxial regime, Sheppard outlined a design procedure for a maximally flat filter using BPM [78]. Wang and Jabbour developed models and design methodologies for binary phase pupils in high NA objectives for the scalar and vectorial case, respectively [79, 110]. The use of BPM in low NA wavefront coding was studied by Ben Eliezer et al. [94, 96]. Equation 3.3 defines the transmission function for a pupil with BPM as,

$$T(w) = \text{circ} \left( \frac{w}{\sin \alpha} \right) \exp \left( ik \pi \sum_{j=0}^{(N_{az}-1)/2} \left[ \theta(w - w_{\text{bpm},2j+1}) - \theta(w - (w_{\text{bpm},2j+1} + dw_{\text{bpm},2j+1})) \right] \right).$$

(3.3)

The annular transmission function is written as the summation of the Heaviside functions, $\theta(w)$. The vector $w_{\text{bpm}}$ denotes the radial position of each $0 - \pi$ phase transition, and is of length $N_{az} + 1$, where $N_{az}$ is the number of annular zones. The first term in the vector, $w_{\text{bpm},1}$, always equals 0, and the last term $w_{\text{bpm},N_{az}}$ always equals the extent of the normalized pupil, $\sin(\alpha)$. The difference in radius between adjacent zones is denoted by $dw_{\text{bpm}}$ which is of length $N_{az}$ by definition.

Figure 3.5(a) shows the PSF and axial intensity for a seven zone BPM filter with the transitions located at $w_{\text{bpm}} = [0, 0.0896, 0.2852, 0.4869, 0.6136, 0.6755, 0.7688, 1] \times \sin(\alpha)$ designed by Wang et al. [110]. The PSF maintains an extended focus in the region between $\pm 1.5$ $\mu$m, and its axial intensity is flat for approximately 1 $\mu$m. The PSF also shows secondary foci outside the maximally flat design region. The depth of field of the PSF could theoretically be extended further and more uniformly by utilizing more binary phase zones. However, with each new phase transition the solution space grows considerably making design more cumbersome, often requiring heuristic optimization algorithms [80]. Minimizing the number of phase transitions needed greatly simplifies the design process.

\[1\] Note that this phase mask was originally designed for a 0.85 NA objective instead of the 1.4 NA objective shown. This scales the profile axially, but does not affect its intensity profile.
Figure 3.4: Binary Phase Modulated Pupil Plane. The mask consists of annular zones which alternate between $\pi$ (gray) and 0 (blue) phase shifts with respect to the central wavelength.

In looking at the MTF, Fig. 3.5(b), high contrast levels in the EDF region can be seen. Unfortunately numerous zeros and contrast reversals are also observed, even at low spatial frequencies. This limits the effectiveness of BPM designs when the image is not sparse, and/or if any digital post-processing steps are required. The number of zeros tends to increase proportionally with the number of phase transitions used in the filter, and are virtually impossible to eliminate. So while BPM offers an impressive amount of control in engineering the PSF, its practical implementation is limited to two- or three- zone filters in hybrid imaging applications. However, as will be shown, BPM can be leveraged in combination with other pupil mask designs to greatly simplify their use and effectiveness in EDF imaging.

3.1.4 Unbalanced OPD Partitioned Pupils

The final pupil mask studied in this comparison is the incoherently partitioned pupil (IPP). It was first proposed and demonstrated by Abrahamsson et al. [89], while the theory was formalized by Chu et al. [86]. A tiered phase mask, as shown in Fig. 3.6, is placed at the pupil plane. If the optical path difference (OPD) of each tier is made to be longer than the coherence length, $\ell_c$, of the light
Figure 3.5: (a) 7-zone Binary Phase Modulated x-z PSF and axial intensity, (b) 7-zone Binary Phase Modulated MTF along $k_x$. 
being imaged, then the pupil is effectively partitioned into the annular sub-pupils, $\Sigma_1, \Sigma_2, \ldots, \Sigma_M$, where the total intensity PSF of the system is simply the summation of the intensity PSFs from each sub-pupil. Partitioning the pupil in this way increases the DOF proportionally to the number of tiers.

As outlined in [86, 88] the intensity PSF for a pupil partitioned by unbalanced OPD illuminated by a broadband source can be written as,

$$h(x, y; z) = \sum_{j=1}^{M} h_{\Sigma_j}(x, y; z) + \text{cross terms},$$

(3.4)

where in the case of unpolarized light,

$$h_{\Sigma_j}(x, y; z) = \int S(\nu) \left( \frac{\nu}{c} \right)^2 \frac{1}{2} (I_H(x, y; z) + I_V(x, y; z)) \, d\nu.$$  

(3.5)

$h_{\Sigma_j}(x, y)$ is the intensity PSF for given subregion, $\Sigma_j$, of the pupil, $\nu$ is the temporal frequency of light, and $S(\nu)$ is the spectrum of the light being imaged. The $\nu^2$ is factored out of the intensity PSFs and can be traced back to the factor of $k$ in Eq. 2.5. The cross terms go to zero when the OPD of each tier is much greater than the coherence length of the signal light [86]. The OPD is found from the refractive index difference of the phase mask, $(n_{pm} - 1)$, times the height of the given tier $t_j$. Assuming each tier is of equal height, the pupil function for a given partition $j$ can
be written as,

$$ T_j(w) = \left( \theta(w - w_{\Sigma,j}) - \theta(w - w_{\Sigma,j+1}) \right) \exp \left( ik(j-1)n_{pm}t \right), \quad (3.6) $$

where \( w_{\Sigma} \) is a vector which contains information on the radial extent of each partition. Like the \( w_{bpm} \) vector its first element is always 0 and its last element represents the extent of the pupil, \( \sin(\alpha) \). The phase term is included in the pupil function to account for the delay of a given partition.

Abrahamsson et al. claimed 100 \( \mu m \) as a sufficient OPD to decorrelate a given region of the pupil in fluorescence microscopy applications, however this was never quantitatively justified [89]. While Abrahamsson’s claim was experimentally validated, the following analysis shows that an OPD significantly less than 100 \( \mu m \) should still be sufficient for pupil partitioning. Fluorescence microscopy utilizes a wide number of protein labels, each with their own spectral properties. One of the most commonly used labels is green fluorescent protein (GFP), as such it was used to estimate the necessary OPD for each tier of the partitioning pupil mask. The peak of the emission spectra for GFP is located at 509 nm with a FWHM of 40nm. The emission spectrum passes through a filter before reaching the camera which, for the purposes of this analysis, assumes the same center wavelength of 509 nm with a constant bandpass of 40nm. Under these assumptions the emission spectrum of GFP can be approximated as a Gaussian curve centered at 509 nm with a FWHM of 17nm when truncated by the emission filter. The coherence length can be defined as the OPD where fringe visibility reduces to 50% of maximum, or in terms of spectral properties as,

$$ \ell_c = c\tau_c = \frac{2\ln(2)}{\pi n} \frac{\lambda_c^2}{\Delta \lambda}, \quad (3.7) $$

where \( c \) is the speed of light in vacuum, \( \tau_c \) is the coherence time, \( \lambda_c \) is the central wavelength of the emission spectra, and \( \Delta \lambda \) is the FWHM of the emission spectra. Using the parameters defined above, the coherence length for GFP in fluorescence microscopy is estimated as 4.43 \( \mu m \). Conservatively estimating that a path difference 10 times longer than the coherence length is needed for complete decorrelation, an OPD of 45 \( \mu m \) should be more than sufficient for pupil partitioning in fluorescence microscopy. Thus Abrahamsson’s mask was many times thicker than needed.

Figure 3.7(a) shows the PSF and axial intensity of a partitioned pupil with 15 annular
Figure 3.7: (a) 15-zone partitioned pupil x-z PSF and axial intensity, (b) Partitioned pupil MTF along $k_x$. 
zones. The radius was scaled so that each zone has equal area. The PSF exhibits both radial and axial symmetry. The PSF is elongated over a range of $\pm 10 \ \mu m$, with the FWHM occurring near $\pm 3.5 \ \mu m$. The axial intensity features a Gaussian type profile through focus. The MTF, Fig. 3.7(b), shows a high degree of uniformity through focus, although it never quite reaches the traditional diffraction limited cutoff frequency. This indicates that the central lobe of the partitioned pupil PSF is slightly broader than the Airy disk. There are no zeros in the transfer function. The modulation depth is worse than the on-axis CPM, but better than the axicon. The partitioned pupil incorporates many of the best features of the aforementioned phase masks. The highly symmetric design and well conditioned MTF indicate partitioned pupils for EDF are a highly viable solution.

<table>
<thead>
<tr>
<th></th>
<th>Cubic</th>
<th>Axicon</th>
<th>BPM</th>
<th>IPP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radial Symmetry</td>
<td>×</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Axial Symmetry</td>
<td>✔</td>
<td>×</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Focus Invariance</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Axial Intensity Uniformity</td>
<td>✔</td>
<td>×</td>
<td>✔</td>
<td>×</td>
</tr>
<tr>
<td>No Zeros in MTF Prior to Cutoff</td>
<td>✔</td>
<td>✔</td>
<td>×</td>
<td>✔</td>
</tr>
</tbody>
</table>

Figure 3.8: Comparison matrix for EDF pupil mask properties. A solid check mark indicates a property is completely satisfied. An outlined check mark indicates a property is partially or conditionally satisfied. An “×” indicates a property is not satisfied.

From the analysis presented in this section, the general properties of EDF phase masks are summarized in Fig. 3.8. Qualitative properties are listed in the first column of the matrix and the pupil masks are indicated in the top row creating a comparison matrix. A solid check mark indicates a given property is completely exhibited by the phase mask. An outlined check mark
indicates a given property is partially or conditionally satisfied, and a “x” indicates a property is not satisfied by the mask.

3.2 Incoherently Partitioned Pupil with BPM

The results of the previous section suggest that incorporating wavefront modulation to an incoherently partitioned pupil would yield an additional degree of freedom in the design of EDF pupil masks. Chu et al. demonstrated that a two tier partitioned pupil mask consisting of logarithmic axicons outperformed both a standard IPP mask as well as an unpartitioned aspheric pupil mask [87, 88]. However, the fabrication of such systems would become increasingly intractable as the number of partitions increased. Thus an incoherently partitioned pupil modulated by binary phase rings (IPP-BPM) is proposed. The IPP-BPM mask would still give added flexibility in the design of EDF systems, and could be readily fabricated using a combination of 3D printing and lithography techniques.\(^2\) In this section the IPP-BPM mask is shown to incorporate the best properties of its two parent pupil masks.

A realization of the proposed phase mask with one binary phase transition per partition can be seen in Fig. 3.9. The black circles in the top view indicate the annular regions that have been incoherently partitioned. Within a given partition the gray region represents a phase delay of \(\pi\) for a given central wavelength \(\lambda_c\), and the blue region represents no phase delay. The thickness of the binary phase modulation which yields a \(\pi\) phase shift is denoted by, \(d_\pi\), and found by calculating the OPD necessary to delay the wavefront by a half-wave.

\[
d_\pi = \frac{\lambda_c}{2(n_{pm} - 1)} \tag{3.8}
\]

The transmission function for a given partition of the IPP-BPM can be written as the product of

\(^2\) Fabrication details are discussed in Chapter 6.
Eqs. 3.3 and 3.6,

\[ T_j(w) = \text{circ}\left(\frac{w}{\sin \alpha}\right) \left(\theta(w - w_{\Sigma,j}) - \theta(w - w_{\Sigma,j+1})\right) \exp\left(ik(j - 1)n_{pm}t\right) \times \exp\left(ik\pi \sum_{p=0}^{[(N_{az}-1)/2]} \left[\theta(w - w_{\text{bpm},2p+1}) - \theta(w - (w_{\text{bpm},2p+1} + dw_{\text{bpm},2p+1}))\right]\right). \] (3.9)

Combining incoherent pupil partitioning and BPM exponentially expands the design space. One could vary the partition radii, the number of binary phase transitions in a given partition, and the radii of those phase transitions. This gives a tremendous amount of control over the design of PSFs and/or MTFs; it also means optimizing over all of these free parameters simultaneously quickly becomes infeasible, even given modern computing power. Thus a series of constraints are laid out in the following subsection to provide some preliminary understanding of the IPP-BPM phase masks and better guide the design and optimization process. A comparison of modulated and unmodulated partitioned pupils is included throughout the analysis. The following design analysis is presented with the goal of better understanding IPP-BPM phase masks for EDF optical microscopy, and is not intended to be all encompassing.
3.3 Design Considerations and Optimization of IPP-BPM Masks

3.3.1 Radii of Pupil Partitions

The first step in constraining the design space is to determine the optimal distribution of radii for each pupil partition. There are two obvious choices for this distribution: use a fixed value for the differential in radius between adjacent partitions or scale the radii so that the exposed region of each partition has equal area. The former shall be referred to as equal radius while latter shall be referred to as equal area. The equal radius distribution scales the partitions linearly with the radius, while the equal area distribution scales the partitions with the square root of the radius. Chu utilized an equal area distribution, although provides no justification [86]. Abrahamsson did not directly specify the distribution of radii in her device, although it appears to also be an equal area distribution based on the figures available [89]. Figure 3.10 plots the FWHM of the axial intensity against the number of partitions for the two radial distributions. Both configurations extend the DOF approximately linearly with the number of annular partitions, although it is not immediately evident from the theory why this is the case. It is clear, however, that equal area distribution achieves the same extension in the DOF with a fewer number of partitions. From linear fits the slopes were found to be 0.465 $\mu$m/partition for the equal area case, and 0.267 $\mu$m/partition for the equal change in radius case. The analysis of Fig. 3.10 suggests an IPP mask with 5 partitions would achieve a $4.8 \times$ improvement in DOF. This value falls in-line with the $4.7 \times$ improvement reported in [89]. These relations can be used as a design tool when optimizing a partitioning mask for a specific extended DOF. Minimizing the number of partitions is critical as it reduces manufacturing cost and complexity. Thus equal area partitioning shows itself as the more desirable configuration. For this reason, the equal change in radius configuration is no longer considered in this thesis. It should be noted that this analysis does not exclude the viability of irregularly distributed partitions.
Figure 3.10: DOF as a function of pupil partitions. The FWHM of the axial intensity is taken as an analog for DOF. The equal area partitioned pupil (blue) extends the DOF at approximately $1.75 \times$ the rate of the equal change in radius partitioning (red). Simulated data points are shown as x’s and o’s respectively along with the best fit lines.

### 3.3.2 Number and Location of Binary Phase Transitions

The next step in constraining the design space is to determine the optimal number of phase transitions per partition. As previously discussed, each phase transition in a BPM mask contributes to the number of zeros and contrast reversals in the MTF. While this effect is mitigated somewhat by the pupil partitioning it still poses a significant limitation on the phase mask if any post-processing is required. This suggests that minimizing the number of binary phase transitions is ideal for EDF hybrid imaging microscopy. Another reason to minimize the number of phase transitions is to simplify fabrication. As more transitions are used the feature sizes of the mask begin to shrink, especially at the periphery of the pupil. This leads to tighter tolerances for the pupil mask and begins to complicate the fabrication process. Thus the designs presented herein are limited to one binary phase transition (two zones) per partition.

The final free variables are the relative positions of the phase transition within each partition. The relative position of a phase transition, $w_{rel}$, within a given sub-pupil is defined as the linear scaling of the minimum and maximum radius of the sub-pupil to the interval $[0,1]$. The
relationship between $w_{\text{bpm}}$ and $w_{\text{rel}}$ for the IPP-BPM mask is given as,

$$w_{\text{bpm},q} = [w_{\Sigma,j} + (w_{\Sigma,j+1} - w_{\Sigma,j})]w_{\text{rel},q},$$

with $q$ phase transitions where the first ‘transition’ always occurs at 0 and the final transition occurs at 1. As an illustrative example, for an IPP-BPM mask with 4 equally spaced annular phase zones the relative location of the phase transitions is denoted by the following 5 element vector, $w_{\text{rel},q} = [0, 0.25, 0.50, 0.75, 1.00]$. Since a central goal of this work is to improve the usefulness of EDF phase masks in quantitative imaging experiments, the relative location of the phase transitions should be chosen to generate a PSF with a maximally flat response. This would allow for an accurate photometric comparison of emitters from different focal planes. If the axial intensity oscillates heavily, as in the case of the axicon, it would be impossible to distinguish between a weak emitter and a strong emitter at different planes.

With this goal in mind two strategies were explored for determining the location of the binary phase transitions. The first emphasizes simplicity and mandates that the transition occurs at the same relative position within each partition. The impetus behind this design choice comes from the empirical observation that for a given relative location of the single binary phase transition, the PSF generated by each sub-pupil follows the same basic form, but is scaled in depth. Thus, the PSF from each partition extends the total PSF in depth without changing its overall axial profile. Under this constraint the relative location of the binary phase transition determines the flatness of the axial intensity response. A design trade-off exists between the flatness and the overall DOF as shown in Fig. 3.11.

The location of the phase transitions was manually tuned to limit fluctuations in the axial intensity to less than 7.5% over the central lobe. The 7.5% ripple point was chosen to balance the overall DOF with the axial flatness for the purposes of the comparisons to follow. The specific flatness constraint will of course be ultimately tied to the given application. The axial intensity FWHM for the two-zone IPP-BPM mask is plotted against the number of partitions in Fig. 3.12 (right axis). The axial intensity FWHM of the IPP mask is also plotted for comparison.
Flatness vs DOF Tradeoff

Figure 3.11: IPP-BPM axial intensity profiles with varying amount of ripple. A trade-off exists between the flatness of the axial intensity response and the axial FWHM. Profiles for 0% (yellow), 5% (red), 10% (green), and 15% (white) ripple shown, to maintain the prescribed flatness the relative location of the phase transition had to be decreased with the increasing number of pupil partitions. This is shown as the dashed line in Fig. 3.12 plotted against the left axis. The IPP-BPM mask increases the DOF at $2.1 \times$ the rate of the unmodulated IPP mask. The best fit line yields a rate of 0.980 $\mu$m/partition.

The 7.5% ripple PSF and MTF generated by an IPP-BPM mask with 7 partitions are shown in Fig. 3.13. The relative phase transition vector is $w_{rel} = [0, 0.837]$. The PSF shows a high degree of symmetry and uniformity in both the axial and transverse planes. The 4 $\mu$m region about $z = 0$ shows a flat response (within the designed ripple constraint) which would be ideal for quantitative intensity comparisons, and has an axial FWHM of 6.85 $\mu$m. The MTF features none of the zeros or contrast reversals prior to the cutoff frequency normally seen in unpartitioned BPM filters. The IPP-BPM mask reproduces the PSF flatness of the BPM pupil while maintaining a well conditioned MTF for further post-processing. While similar results have been achieved using radially polarized beams and binary phase structures [111]. IPP-BPM represents the realization of this concept in conditions where one does not have complete control over the structure of the light being imaged.

The second strategy for determining the position of the relative binary phase transitions was to use stochastic optimization. Specifically, particle swarm optimization [112] was used to search
Figure 3.12: DOF as a function of equal area pupil partitions with and without BPM (right axis). The BPM pupil (red diamonds) extends the DOF at approximately 2.1\times the rate of the unmodulated pupil (blue squares). The relative position for each phase transition (red diamonds w/ dashed line) decreases with increasing number of partitions (left axis).

the design space. As with the design previously shown, the IPP-BPM mask was still limited to one phase transition (two phase zones) per partition and 7 pupil partitions. The optimizer’s merit function was defined as the Euclidean distance between the axial intensity profile of an ideal EDF PSF and the axial intensity profile of a given PSF generated by the particle swarm optimizer. The ideal EDF axial intensity profile was defined as having unit relative intensity over a region $\pm 3.5 \mu m$ from the traditional focal plane, and 0 intensity outside of this region. The starting point for the optimization was the 7.5% ripple design discussed above. The best result from a particle swarm optimization consisting of 50 particles running for 250 iterations is shown in Fig. 3.14. The relative location of each partition’s phase transition is $w_{rel,2} = [0.3005, 0.3426, 0.9801, 0.9111, 0.7289, 0.7215, 0.5567]$.

The resulting PSF is very similar to Fig. 3.13(a) showing a maximum intensity ripple of 7.67% and axial intensity FWHM of 6.82 $\mu m$. However, the through focus PSF shows more defined side lobes, as opposed to the more monotonic intensity decay of the previous design. Additionally, zeros can be seen in the MTF at high spatial frequencies, and the overall cutoff frequency is lower. These negative properties could potentially be mitigated by expanding or redefining the merit function to include the MTF. Expanding the optimizer in this way creates a new design
Figure 3.13: PSF (a) and MTF (b) for an IPP-BPM pupil with 7 partitions and 2 binary phase zones per partition. The relative location of the phase transition occurs at $w_{rel,2} = 0.837$. 
Figure 3.14: PSF (a) and MTF (b) for an IPP-BPM pupil with 7 partitions and 2 binary phase zones per partition designed via particle swarm optimization.
problem in determining the optimal terms and weights for the merit function as well as significantly increasing computational cost. Due to limited resources these modified merit functions were not pursued, however they will be the subject of future investigations as there is undoubtedly room for improvement. Due to the current frequency space deficiencies of the stochastically optimized pupil mask, the 7.5% axial intensity ripple design with constant relative locations of the binary phase transitions will be used in all further analysis.

3.4 Off-axis Performance of the IPP-BPM Mask

The off-axis performance of the IPP-BPM was studied in order to assess the effect on the PSF. Assuming a $1024 \times 1024$ region of interest and a pixel pitch of 6.5 $\mu$m on the focal plane, the full field of view (FOV) for a $63 \times$ objective is slightly over 105 $\mu$m. Emitters were simulated at 0 $\mu$m, 25 $\mu$m, and 50 $\mu$m off of the optics axis at 45° so as to cover a significant portion of this FOV. Zoomed in images of transverse IPP-BPM PSFs at the 3 field points at the axial planes of +4 $\mu$m, 0, and −4 $\mu$m are shown in Fig. 3.15. While the central lobe of the off-axis PSFs show some elongation along the direction of displacement, the effect is minimal.

Note that the above simulation does not fully encompass all of the effects of the true physical structure of the IPP-BPM pupil mask when implemented in practice, and thus should only be considered as an initial approximation. Specifically for thick pupil partitioning masks, one would expect to observe scattering effects from off-axis light propagating through the structure. These adverse effects can be mitigated to some extent in an experimental setup by placing the partitioning mask at a conjugate pupil plane through the use of low NA relay optics. The relatively small focusing angle of the relay optics would minimize scattering from off-axis light.

This concludes the qualitative comparison and analysis of EDF phase masks. Despite various design constraints it has been shown that the IPP-BPM mask exhibits the best properties of its parent masks: axial and radial symmetry, a maximally flat response, and no zeros in the MTF prior to the cutoff frequency. Additionally, it achieves the same DOF performance as an IPP mask with fewer partitions. There is potential for improved designs through continued use
Figure 3.15: Off-axis PSFs of IPP-BPM pupil mask
of stochastic optimization techniques, and extending the DOF represents only one possible use for this new family of optical elements. By exploring different imaging tasks and design constraints one could envision a multitude of applications for IPP-BPM masks. A quantitative comparison of the various EDF pupil masks now follows.

### 3.5 Comparison Metrics EDF Pupil Masks

The following metrics will be utilized to evaluate the quantitative effectiveness of the IPP-BPM pupil mask presented in the previous section to the EDF pupil masks in the literature: axial Strehl ratio (SR), average transverse PSF width, and PSF Hilbert space angle. A description and definition of each metric in the context of fluorescence microscopy is provided, followed by a comparison study.

#### 3.5.1 Axial Strehl Ratio

The first metric used is the axial Strehl ratio (SR) [79]. The axial SR is defined as the peak intensity of a given aberrated (or in this case EDF modulated) PSF over the peak intensity of a diffraction limited PSF. Obviously the higher the SR the better the SNR will be in the recorded image. For weakly fluorescing emitters it is essential to maintain the highest axial SR possible to ensure sufficient signal for post-processing. Since the quantity of interest is the SR over an axial extent it is defined as a function of defocus,

$$SR(z_d) = \frac{h(x = 0, y = 0, z_d)}{h_{DL}(x = 0, y = 0, z_d = 0)},$$

where $h$ is the intensity PSF for a given extended depth of field pupil, and $h_{DL}$ is the peak intensity of a diffraction limited PSF. For the case of the CPM the numerator in Eq. 3.11 is replaced by the maximum intensity projection in order to make a fair comparison.

The SR yields three metrics to evaluate EDF systems: its FWHM is a measure of DOF, its mean value over that DOF, $SR_\mu(DOF)$, is a measure of the relative signal strength, and its full width at 90% maximum (FW-90%M) is a measure of PSF flatness.
3.5.2 Transverse PSF FWHM

The transverse resolution of an optical system can be characterized by the central lobe width of the PSF. This is an especially pertinent metric when working in the spatial domain as is the case in localization microscopy [113, 114, 115]. The resolution of any biological system consisting of sparse, point-like emitters can also be readily described in this way. For the purposes of this thesis, the transverse resolution will be defined as the FWHM along the x axis of the average transverse PSF over the DOF. This quantity will be denoted as $FWHM_t$ in the comparisons to follow. The exception to this is the clear aperture, diffraction limited PSF which is simply taken at the plane of best focus to serve as a benchmark.

3.5.3 PSF Hilbert Space Angle

One of the main benefits of using pupil function engineering is that a single deconvolution filter is used to restore contrast over the entire extended DOF region. The effectiveness of this post-processing step is dependent on focus invariance of the PSF/MTF. The focus invariance can be quantified by the angle of the PSF through focus in Hilbert space, $\theta_H$ [56]. Defined as,

$$\cos \theta_H(z_d) = \frac{\langle h(x,y,0), h(x,y,z_d) \rangle}{\|h(x,y,0)\| \|h(x,y,z_d)\|},$$  \hspace{1cm} (3.12)

the angle gives a metric by which one can measure the difference in the PSF through focus compared to the PSF at the best focus plane. For pupil mask designs which shift the plane of best focus (i.e. axicons and their derivatives), $h(x,y,0)$, is taken to be the plane at the center of their axial FWHM. The inner product of the best focus and defocused PSFs is defined in the continuous domain as,

$$\langle h(x,y,0), h(x,y,z_d) \rangle = \int_{-\infty}^{\infty} h(x,y,0)h(x,y,z_d) \, dx \, dy. $$  \hspace{1cm} (3.13)

The norms in Eq. 3.12 are defined in the continuous domain as,

$$\|h(x,y,0)\| = \left[ \int_{-\infty}^{\infty} h(x,y,0)h(x,y,0) \, dx \, dy \right]^{\frac{1}{2}}, $$ \hspace{1cm} (3.14)

and,

$$\|h(x,y,z_d)\| = \left[ \int_{-\infty}^{\infty} h(x,y,z_d)h(x,y,z_d) \, dx \, dy \right]^{\frac{1}{2}}. $$  \hspace{1cm} (3.15)
The discrete equivalents of the Eqs. 3.13 - 3.15 are used in the numerical analysis to follow.

3.6 Quantitative Analysis of Increasing Pupil Partitions

The analysis accompanying Figs. 3.10 and 3.12 began to explore the effect of increasing pupil partitions on the axial FWHM in both the unmodulated and binary phase modulated cases. This will now be extended to include a more quantitative assessment of both PSFs using the metrics described in the previous section.

Tables 3.2 and 3.3 show the three axial SR metrics along with the transverse FWHM for both types of partitioned pupils. The number of partitions is denoted by the number following the IPP abbreviation. For the IPP-BPM pupils the number following the BPM abbreviation denotes the number of annular phase zones per partition. Due to the different rates in axial FWHM/partition the unmodulated pupils range from a clear aperture (1 partition) to 20 partitions, while the BPM modulated pupils only span from 1-7 partitions. The PSFs were calculated over an axial extent of ±4 µm. All other parameters are identical to the values listed in Table 3.1. The average axial SR is taken only over the axial FWHM for the given PSF, except for the case of 20 pupil partitions which is taken only over 8 µm since its axial FWHM extends beyond the simulation range.

Table 3.2: PSF comparison metrics for IPP EDF phase mask.

<table>
<thead>
<tr>
<th>Pupil Mask</th>
<th>DOF [µm]</th>
<th>SRµ(DOF)</th>
<th>FW-90%M [µm]</th>
<th>FWHMt [µm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear Aperture</td>
<td>0.478</td>
<td>0.831</td>
<td>0.190</td>
<td>0.211</td>
</tr>
<tr>
<td>IPP02</td>
<td>0.934</td>
<td>0.406</td>
<td>0.367</td>
<td>0.234</td>
</tr>
<tr>
<td>IPP05</td>
<td>2.320</td>
<td>0.162</td>
<td>0.898</td>
<td>0.241</td>
</tr>
<tr>
<td>IPP10</td>
<td>4.638</td>
<td>0.081</td>
<td>1.787</td>
<td>0.242</td>
</tr>
<tr>
<td>IPP15</td>
<td>6.956</td>
<td>0.054</td>
<td>2.678</td>
<td>0.242</td>
</tr>
<tr>
<td>IPP20</td>
<td>9.316</td>
<td>0.043</td>
<td>3.570</td>
<td>0.242</td>
</tr>
</tbody>
</table>

As expected, SRµ decreases with the increasing DOF for both pupil types. For an equivalent average SR the unmodulated pupils show a longer DOF. Conversely, the BPM modulated pupils exhibit a longer FW-90%M for an equivalent average SR. See IPP15 and IPP06-BPM02 as an approximate example of both cases. These results fall in line with expectations as the IPP-BPM...
Table 3.3: PSF comparison metrics for IPP-BPM EDF phase mask.

<table>
<thead>
<tr>
<th>Pupil Mask</th>
<th>DOF [µm]</th>
<th>SRₚₙ (DOF)</th>
<th>FW-90% M [µm]</th>
<th>FWHMᵣ [µm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPP01-BPM02</td>
<td>0.926</td>
<td>0.435</td>
<td>0.629</td>
<td>0.235</td>
</tr>
<tr>
<td>IPP02-BPM02</td>
<td>1.948</td>
<td>0.183</td>
<td>1.230</td>
<td>0.251</td>
</tr>
<tr>
<td>IPP03-BPM02</td>
<td>2.962</td>
<td>0.117</td>
<td>1.805</td>
<td>0.248</td>
</tr>
<tr>
<td>IPP04-BPM02</td>
<td>4.037</td>
<td>0.086</td>
<td>2.441</td>
<td>0.245</td>
</tr>
<tr>
<td>IPP05-BPM02</td>
<td>4.923</td>
<td>0.070</td>
<td>2.922</td>
<td>0.244</td>
</tr>
<tr>
<td>IPP06-BPM02</td>
<td>5.879</td>
<td>0.058</td>
<td>3.473</td>
<td>0.243</td>
</tr>
<tr>
<td>IPP07-BPM02</td>
<td>6.849</td>
<td>0.050</td>
<td>4.046</td>
<td>0.242</td>
</tr>
</tbody>
</table>

Pupils are designed to have a flatter PSF response than the unmodulated pupils. The cost for achieving this response is a slightly lower signal strength. For the unmodulated pupils the DOF is consistently $\sim 2.6 \times$ longer than the FW-90% M regardless of the number of partitions. The factor is approximately $1.7 \times$ for the IPP-BPM masks, and again holds relatively constant with respect to the number of partitions. This suggests that the IPP-BPM mask would be more useful in applications where one wishes to limit the background contributions of features beyond targeted DOF region of interest. For a visual comparison, axial intensity profiles are shown in Fig. 3.16 with IPP masks on the left and IPP-BPM masks on the right. It is observed that the unmodulated pupils begin to stabilize after 15-20 partitions, and the IPP-BPM masks stabilize around 5-7 partitions. The axial behavior of the PSFs follow largely as one would expect with the increasing number of partitions.

Somewhat less obvious is the behavior of the average transverse PSF over the DOF. The transverse FWHM begins broadening with an increasing number of partitions, as one would expect, but quickly stabilizes. It settles at a value of 242 nm in the unmodulated pupils. Oddly, the IPP-BPM transverse PSF is most broad after the first partition, then begins to narrow towards the same 242 nm value. Both types of partitioned transverse PSFs begin to broaden significantly when their normalized intensity drops below 30%. At these intensity values the PSF starts to broaden with the increasing number of partitions as one would expect, Fig. 3.17.

Figure 3.18 shows the PSF Hilbert space angle through focus for both pupil types. As
the number of pupil partitions increase the PSFs become more uniform through focus. The unmodulated pupils show more focus invariance than the BPM pupils. For the IPP mask there is little difference between the 15 and 20 partition cases, and 10 partitions would be suitable in most applications that require post-processing. The IPP-BPM mask stabilizes after 5 partitions.
3.7 Quantitative Comparison of EDF Pupils

The partitioned pupils discussed in the previous section are now quantitatively compared against a clear aperture, an axicon, and the cubic phase mask. The simulation region is again ±4\(\mu\)m. The unpartitioned BPM pupil mask [110] has been omitted as its usable EDF range is significantly less than the other phase masks, and a comparable design is not readily available in the literature. Note that the data for the axicon has been shifted to the negative by 7.75\(\mu\)m to account for the focusing power of the phase mask. The magnitude of the peak-valley phase delay for the cubic and axicon are 25.8\(\lambda\) and 21.5\(\lambda\) respectively at the central wavelength. They are compared along side the IPP15 and IPP07-BPM2 partitioned pupils. The strengths of all of the phase masks are set so that the DOF is roughly equivalent.

Table 3.4 shows the PSF metrics for all of the phase masks. The strength of each mask has been set so they all have a roughly equivalent DOF.\(^3\) The IPP15 mask shows the highest average

\(^3\) Note the DOF of the cubic PSF is somewhat harder to determine as its intensity does not decay regularly with defocus. As such, its DOF was estimated as the planes where the transverse PSF does not diverge significantly from the best focus PSF.
Table 3.4: PSF comparison metrics for a variety of EDF phase masks with roughly equivalent axial FWHM.

<table>
<thead>
<tr>
<th>Pupil Mask</th>
<th>DOF [$\mu$m]</th>
<th>SR$_a$(DOF)</th>
<th>FW-90%M [$\mu$m]</th>
<th>FWHM$_t$ [$\mu$m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clear Aperture</td>
<td>0.478</td>
<td>0.831</td>
<td>0.190</td>
<td>0.211</td>
</tr>
<tr>
<td>Axicon</td>
<td>6.952</td>
<td>0.039</td>
<td>0.655</td>
<td>0.203</td>
</tr>
<tr>
<td>Cubic</td>
<td>6.950</td>
<td>0.009</td>
<td>6.454</td>
<td>0.604</td>
</tr>
<tr>
<td>IPP15</td>
<td>6.956</td>
<td>0.054</td>
<td>2.678</td>
<td>0.242</td>
</tr>
<tr>
<td>IPP07-BPM02</td>
<td>6.849</td>
<td>0.050</td>
<td>4.046</td>
<td>0.242</td>
</tr>
</tbody>
</table>

Axial SR at 5.4%, with the IPP-BPM mask just behind it. The cubic PSF has the weakest average signal strength by far which makes sense given that its PSF is spread over significantly more area in the transverse plane than the others. The cubic does have the flattest intensity response, as seen in Fig. 3.19. Although one must keep in mind that this is the maximum intensity projection of the cubic, and not its true axial SR. Of the radially symmetric PSFs, the IPP-BPM mask shows the most flatness. In terms of the transverse PSF width the axicon creates the narrowest central lobe, even when compared to the diffraction limited PSF. This is a known property of the quasi-Bessel beam produced by axicons and similar phase masks. The cubic again performs considerably worse than the other masks in this category creating an average central lobe that is 2.5-3× wider than the other PSFs.\(^4\) The transverse PSF profiles can be seen in Fig. 3.20. From this comparison the following claims can be made: the IPP mask provides the best signal strength for the given DOF, the IPP-BPM mask provides the most flatness of the radially symmetric phase masks, and the axicon provides the narrowest transverse PSF that can even be considered super resolving in some applications.

The Hilbert space angles of the PSFs through focus are plotted in Fig. 3.21. The clear aperture and cubic PSFs show the worst focus invariance. The radially symmetric EDF PSFs are all grouped together with none of them exceeding an angle of 0.215 radians. The IPP mask shows the best performance followed by the IPP-BPM mask and the linear axicon. These results

\(^4\) The cubic was designed to be used in conjunction with post-processing. This comparison is carried out in Chapter 5 and is more relevant to the CPM than the pre-processing analysis shown here.
Figure 3.19: Axial PSFs through focus for various phase masks and a clear aperture.

Figure 3.20: Average transverse PSFs through focus for various phase masks and a clear aperture. The legend is the same as in Fig. 3.19.

indicate that post-processed EDF images associated with the radially symmetric PSF will exhibit significantly less artifacts than seen in cubic EDF images.
Figure 3.21: Hilbert space angle comparison of EDF PSFs from various phase masks. The legend is the same as in Fig. 3.19.

3.8 Conclusions

A number of conclusions can be drawn from the analysis presented in this chapter. From the qualitative comparison, general properties of the axicon, cubic, binary phase, and incoherently partitioned pupil phase masks were identified. This investigation indicated that aspects of both the BPM and IPP masks would be ideal for quantitative applications of fluorescence microscopy namely: radial symmetry, axial symmetry, axial intensity uniformity, and no zeros in the MTF prior to the cutoff frequency. Thus a new phase mask was proposed which combined incoherent pupil partitioning with binary phase modulation. The IPP and a maximally flat implementation of the IPP-BPM phase masks were quantitatively analyzed to study the effect of increasing the number of partitions. It was found that the IPP-BPM mask increases the DOF at over twice the rate of the IPP mask while maintaining a maximally flat axial PSF response. It was also found that IPP-BPM achieves its maximally flat response without introducing zeros into the MTF commonly seen with just BPM.

The partitioned phase masks were then qualitatively compared to the axicon and cubic
phase mask. It was found that pupil partitioning to extend the DOF provides the most balance between all of the metrics described above. The quantitative analysis showed that the experimental conditions of a given imaging experiment can have a tremendous impact on effectiveness of the system. Parameters such as the diffraction limited signal strength, desired DOF, desired resolution, and the sparsity of the image must all be considered to create optimal EDF imaging conditions. Additionally, the nature (qualitative or quantitative) and domain (spatial or spectral) of the measurements will have a significant impact on the optimal pupil mask for a given experiment.
The primary goal of this work is the application of the PSF engineering techniques described in the previous chapter to biological microscopy systems. As such it is necessary to do a full treatment of noise sources commonly observed in biological imaging. This is especially important if one wishes to maintain the photometric accuracy of an EDF image post-processing. The EDF literature has primarily focused on non-microscopy applications and thus made many simplifying assumptions when constructing noise and deconvolution models. In addition to the intrinsic photon noise present in imaging, this chapter provides a full treatment of the electronic noise for the three most common focal plane arrays used in biological microscopy: the CCD, the EMCCD, and the sCMOS camera. From these models individual likelihood functions are constructed to tailor the deconvolution problems presented in Chapter 5 to the specific detector being used. Note that in order to maintain a notation convention consistent with the signal processing literature some variables defined in Chapters 1-3 have been reused in Chapters 4 and 5 (e.g. k is used to denote the step of a given iteration in the deconvolution problem as opposed to the wavenumber). This is only done in instances where the context is overtly clear, and the variables are expressly redefined to avoid any confusion.

4.1 Shot Noise

The random fluctuations in photon arrival time is referred to as shot noise [116] and follows a Poisson distribution. The ground truth spatial intensity distribution of the object being observed
is denoted by \( O_0 \) in units of photons. Those photons are collected by the imaging optics and blurred by the intensity PSF, \( h \). The blurred photon distribution is denoted by the convolution \( h * O_0 \). Those photons are then binned into pixels over some exposure time, \( t_e \). This process is denoted by the following sampling function [117],

\[
G_{mn}(O_0) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-Y/2}^{Y/2} \int_{-X/2}^{X/2} O_0(x, y) h(x' - x, y' - y) \times p_s(x' - m\Delta_x, y' - n\Delta_y) dx' dy' dx dy,
\]

(4.1)

where \( p_s \) is the pixel sampling function, \( \Delta_x \) and \( \Delta_y \) are the pixel pitches in \( x \) and \( y \) respectively, \((m, n)\) is a given pixel in the focal plane array, and \( X \) and \( Y \) denote the physical width and height of the focal plane array. Assuming rectangular pixels with a 100% fill factor the sampling function can be represented as,

\[
p_s(x, y) = \text{rect}(x/\Delta_x) \text{rect}(y/\Delta_y).
\]

(4.2)

The probability that \( N_{\text{inc}} \) photons are incident at a given element, \( i \), in the detector plane from an object, \( O_0 \), is given by,

\[
p(N_{\text{inc}}|O_0) = \frac{e^{-G_i(O_0)}(G_i(O_0))^{N_{\text{inc}}}}{N_{\text{inc}}!}.
\]

(4.3)

Note that the two dimensional pixel indicies \((m, n)\) has been replaced by the single index \( i \) for compactness. Shot noise is an inherent property of photon emission and thus a factor in any optical imaging system.

### 4.2 Electronic Detectors

For scientific imaging, particularly biological microscopy and astronomy, there are three types of arrayed detectors that are commonly used: the charged coupled device (CCD), the scientific complementary metal-on-oxide semiconductor (sCMOS), and the electron multiplied CCD (EMCCD). The choice of camera has significant implications on an experiment’s noise, temporal resolution, and SNR. Thus it is essential to select a camera that will yield optimal results for a given set of
experimental conditions. A brief description of each camera’s electronic architecture, noise sources, pros, and cons are provided below. A detailed analysis of each noise term is provided in the sections that follow.

4.2.1 CCD

The CCD camera has been one of the most widely used detectors in biological microscopy over the past 30+ years [118]. Photons captured by the imaging system are focused onto the detector array shown in Fig. 4.1. The photons are converted into electrons at some quantum efficiency (QE), $q$. The generated electrons are then shifted down to a masked storage array. This is done so that another exposure can begin while the previous exposure is being read out the detector electronics. As the electrons are shifted to the storage array spurious electrons are generated in the form of dark current and clock-induced charge (CIC), both of which are described in detail below. The electrons are then transferred to the shift register beneath the storage array again accumulating some spurious charge. Finally, the electrons are converted into a digital signal at some gain factor $g$ by the analog-to-digital readout amplifier. The intensity distribution is now in units of image counts. This process adds what is commonly referred to as readout noise which is also described in the following sections.

4.2.2 sCMOS

Recently the sCMOS camera [119] has been gaining traction in the biological imaging community. It offers faster readout speeds and lower readout noise than traditional CCD cameras. The sCMOS electronic architecture, shown in Fig. 4.2, is significantly different from the CCD. Each pixel has its own readout amplifier which allows for a higher frame rate. The effect of this is that the readout gain factor must now be considered a pixel dependent quantity, $g_i$, as opposed to it being a global quantity in the case of the CCD.

The other major difference in the sCMOS architecture is the electronic shutter, of which there are two types: global and rolling. In the global shutter all of the pixels in the array are
Figure 4.1: Diagram of CCD detector electronics.

Figure 4.2: Diagram of sCMOS detector electronics.
readout simultaneously. In the rolling shutter, which offers lower noise, the array is divided in half horizontally and the rows are read from the center out towards both the top and bottom. There is a 10 μs delay between each row being readout, which would cause a distorting effect for fast moving objects. Generation I sCMOS cameras only feature the global shutter, while generation II models feature either just the rolling shutter, or in some models both.

4.2.3 EMCCD

The final type of arrayed detector commonly used in fluorescence microscopy is the electron multiplying CCD (EMCCD) [120, 121]. EMCCDs offer the highest sensitivity and quantum efficiency of the cameras described with the capability of detecting single photon events. The architecture is essentially the same as a standard CCD, but with an electron multiplication (EM) register in between the shift register and the output amplifier, as shown in Fig. 4.3. The EM register is clocked at a significantly higher voltage increasing the probability of impact ionization. Thus for a given clocking voltage a single electron entering the register will on average be multiplied by the gain factor, $M$. This means that signals which would normally be lost in the readout noise can be amplified to be significantly greater than the noise floor. Typical values for $M$ range from 1 - 1000.

The increased sensitivity comes at a cost however. An undesired effect of the EM register is that it doubles the variance of the amplified signal. This additional noise is commonly referred to as the excess noise factor (ENF). In terms of practical applications this means that EM gain is only effective at boosting the SNR for extremely low photon counts. A value commonly cited in the literature is $10 \, e^-$ [122], however it may range up to $100 \, e^-$ depending on the other camera parameters. Other drawbacks to EMCCDs are that they tend to have slower readout speeds, larger pixel sizes, and are 2-3× more expensive than the latest generation of sCMOS cameras. These limitations mean that EMCCDs are really only beneficial in the case of extremely weak signals.
Figure 4.3: Diagram of EMCCD detector electronics.
4.3 Quantum Efficiency

The quantum efficiency (QE), $q$, is the percentage of incident photons, $N_{inc}$, that get converted into electrons, $N_e$, by the detector, that is $N_e = qN_{inc}$. It is a function of the wavelength of the detector, which in most cases is silicon based. Typical QE curves for the three camera types are shown in Fig. 4.4. The CCD and sCMOS cameras both operate at approximately 60% QE in the middle of the visible spectrum. The EMCCD however is just over 90% QE for the same wavelengths. While the QE can vary slightly over the area of the array it will be considered a fixed quantity for the analysis to follow.

4.4 Fixed Pattern Noise

Even under perfectly uniform illumination (or even absent illumination) detectors will exhibit some spatial variation in their output referred to as fixed pattern noise (FPN) [118]. This noise term arises from inconsistencies in the electronic circuitry of the camera. An example of this are variations in the layer thickness of the semiconductor that arise during the manufacturing process which causes the QE to be spatially dependent. The nature of the FPN is also highly dependent upon the electronic architecture of the detector. It generally appears to be randomly distributed in CCD detectors, whereas column artifacts are more prominent in sCMOS cameras. FPN can be broken down into a constant offset component known as dark signal non-uniformity (DSNU), and signal dependent gain component known as photon response non-uniformity (PRNU).

The DSNU is characterized as the pixel dependent difference from a flat field for a given set of camera parameters assuming no incident illumination. In CCD cameras this is caused by dark current (discussed below), and is generally minimal in state-of-the-art cameras due to their cooling systems. In sCMOS cameras pixel and column variations in the offset and threshold voltages create additional non-uniformities. Regardless of detector the DSNU can be corrected by averaging a series of dark field images for a given set of camera parameters, and subtracting out the bias [123].
Figure 4.4: Quantum efficiency curves from Andor Technology for typical (a) CCD, (b) sCMOS, and (c) EMCCD cameras.
The PRNU can be caused by different pixel areas, the aforementioned variations in QE, and/or variations in the column amplifiers (for sCMOS cameras). This manifests as hot pixels in CCDs, and both hot pixels and hot columns in sCMOS cameras. Pixel and column dependent gain for a given sensor can be estimated by taking a series of images at increasing intensities. The measured intensity values can then be scaled to correct for any variations present. It is essential to have a highly uniform light source to accurately make this measurement. The PRNU for the entire detector is traditionally represented as the standard deviation from the mean response across all pixels in the detector, with ideal values being <1% of the mean signal.

4.5 Spurious Charge

Spurious electrons can be generated as charge is moved through an array and from thermal generation. It comes in two forms: dark current and clock-induced charge (CIC). Both sources follow Poisson statistics and can thus be added to the shot noise term when calculating the SNR.

Dark current, also called thermal or Johnson noise, refers to the electrons generated spontaneously within the imaging array even when no light is incident upon the detector [118]. It is denoted as $c_{dark}$ with units of electrons/pixel. The level of this current is exponentially related to the temperature of the detector. To mitigate this problem, modern scientific cameras are cooled to temperatures as low as -100°C. Many detectors quote the dark current as hundredths of an electron per pixel per second. Thus this form of noise is largely negligible except in only the most sensitive of experiments. Dark current, being a time dependent quantity, is also mitigated by the relatively short exposure times of most biological imaging experiments.

Clocking in a CCD camera refers to the transfer of charge through the detector as an image is being read. As the signal electrons are being transferred there is some probability that spurious electrons can be generated which are referred to as CIC [118]. It is denoted as $c_{CIC}$ with units of electrons/pixel. In a standard CCD the CIC contributes minimally to the overall noise term. In an EMCCD however, where the clocking voltages in the multiplication register are significantly higher, CIC can drastically degrade the image quality. Modern EMCCD cameras have been designed to
minimize this effect, although it is still essential to quantify it for extremely precise photon counting experiments [123].

The total spurious charge can thus be written as, \( c_{sp} = \dot{c}_{dark} t_e + CIC \), where \( t_e \) is the exposure time and \( \dot{c}_{dark} \) is the time derivative of the dark noise in units of electrons/pixel·s. The expected number of electrons generated at a given pixel can then be represented as \( E[N_e] = q \cdot G_t(O_0) + c_{sp,i} \).

### 4.6 Readout Noise

The electrons are then transferred to the readout register and pass through an analog to digital (A/D) output amplifier. The A/D conversion process is a source of additive noise that follows a zero mean Gaussian distribution [118] with a standard deviation, \( \sigma \). For a given number of electrons, \( N_e \), entering the A/D converter the number of image counts, \( f \), is given by the following probability distribution,

\[
p(f|N_e) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}|f-gN_e|^2},
\]

where \( g \) is the A/D gain factor (in units of image counts/electron). In CCD cameras the A/D gain factor and the noise standard deviation are constant parameters, whereas they are pixel dependent quantities in sCMOS cameras. Mean values for \( \sigma \) in CCD, EMCCD, and sCMOS cameras from Andor Technology are quoted 4 \( e^- \), 1.3 \( e^- \), and 20 \( e^- \) respectively. These values are typical cameras used in microscopy, but can vary greatly depending on the sensor. While the read noise is comparably high in the EMCCCD camera, it can be reduced to \(< 1 e^- \) by the multiplication register as discussed in detail in the following section. The latest generation of scientific imaging cameras has largely mitigated the effect of readout noise for moderate or high signal strengths, and can often be omitted from the noise model. For weak signals it must still be taken into account. Additionally, as the readout speed increases so does the noise variance. This can become an additional challenge when imaging at high temporal resolutions.

Other noise components of the analog to digital conversion process are quantization noise
and reset noise [118]. The quantization noise is given by,

\[ N_q = \frac{\text{FWC}}{2^n_{\text{bits}} \sqrt{12}}, \]  

(4.5)

where FWC is the full well count of a pixel in electrons and \( n_{\text{bits}} \) is the bit depth of the camera. Given the well depth and bit depth of modern scientific cameras this term is typically less than 1 e\(^-\) and is thus considered negligible. Reset noise, which is commonly referred to as kTC noise, is caused by uncertainty in the reference voltage of the camera’s sense capacitor when it is reset after measuring each pixel’s charge packet. Typically this would lead to 50 e\(^-\) of noise, however modern cameras contain a correlated double sampler circuit that measures the difference between the reset voltage and the signal voltage for each pixel. This eliminates the need to reset the sense capacitor voltage to the same value for each pixel readout, and mitigates the reset noise.

### 4.7 Electron Multiplication Register Noise

The last source of noise discussed is the multiplication register noise specific to EMCCD cameras. The output probability distribution of an electron multiplying device for a single input electron can be approximated by an exponential distribution [124]. The output distribution for multiple electrons can then be found by convolving \( N_{ie} \) exponential distributions, where \( N_{ie} \) denotes the number of electrons entering into the EM register. Thus for a moderately large mean number of electrons entering the multiplication register, Fig. 4.3, the distribution of output electrons can be well modeled by the following gamma distribution,

\[ \gamma(N_{oe}|N_{ie}) = N_{oe}^{-1} \exp\left(-\frac{N_{oe}}{M}\right) \frac{\Gamma(N_{ie})}{M^{N_{ie}}}, \]  

(4.6)

where \( M \) denotes EM register gain and \( N_{oe} \) denotes the number of electrons exiting from the EM register. This approximation greatly simplifies computation of the output probability distribution, and under normal operating conditions for the EM register was found to be sufficiently accurate [123].
4.8 Likelihood Functions

All of the sources of noise can now be incorporated in order to build a single probability distribution, or likelihood function, to describe the imaging process. Two distinct models, and subsequent simplifying approximations, will be necessary: one to model imaging with CCD and sCMOS cameras, and another for EMCCD imaging. The first step in both models is to combine the effects of shot noise, QE, and spurious charge in a single Poisson distribution. This is achieved through a convolution of a Poisson and binomial distribution in order to combine the effects of shot noise and QE. The resulting distribution is then convolved with another Poisson distribution to account for the effects of spurious charge resulting in,

\[
p(N_e|O_0) = e^{-(q \cdot G_i(O_0) + c_{sp,i})} \frac{(q \cdot G_i(O_0) + c_{sp,i})^{N_e}}{N_e!}.
\] (4.7)

For standard CCD and sCMOS cameras, the next step is to incorporate the readout noise. This is done by convolving the Poisson distribution of Eq. 4.7 with the Gaussian distribution of Eq. 4.4,

\[
p(f_i|O_0) = \sum_{N_e=0}^{+\infty} \frac{e^{-g_i \cdot (q \cdot G_i(O_0) + c_{sp,i})} (g_i \cdot (q \cdot G_i(O_0) + c_{sp,i}))^{N_e}}{N_e!} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2} |f_i - g_i N_e|^2}.
\] (4.8)

While Eq. 4.8 represents the most accurate model for the noise process, it can be cumbersome and computationally slow to work with. When not operating in the ultra weak signal regime one can utilize one of two simplifying approximations to speed up deconvolution without greatly sacrificing photometric accuracy. The first approach is to model the entire process as a non-stationary Gaussian distribution as done in [125, 126]. The other approach, which is used herein, is to model the entire process as a Poisson distribution [127]. Both approximations are similar and rely on the observation that a Poisson distribution with a sufficiently high mean value, \( \mu \), asymptotically approaches a Gaussian distribution of mean, \( \mu \), and variance, \( \mu \). For this case the detector readout noise is approximated as a Poisson distribution with mean \( (g\sigma)^2 \) in units of image counts. Thus one arrives at the following Poisson approximations to the noise model for CCD and
sCMOS cameras,
\[ p(f_i | O_0) = \frac{e^{g_i \cdot (q \cdot G_i(0) + c_{sp,i})} + (g_i \sigma_i)^2 (g \cdot (q \cdot G_i(0) + c_{sp,i}) + (g_i \sigma_i)^2 f_i)}{f_i!}. \]  (4.9)

As previously stated the only difference between the two camera models is whether the gain and readout noise are pixel dependent or global quantities.

A similar approach is taken for EMCCD cameras. Starting again from Eq. 4.7, the EM register noise is incorporated into the model via a composition of the Poisson and gamma distributions,
\[ p(f_i | O_0) = \sum_{N_{ie}=1}^{\frac{f_i}{g}} e^{-g_i \cdot (q \cdot G_i(0) + c_{sp,i})} (g_i \cdot (q \cdot G_i(0) + c_{sp,i}))^{N_{ie}} \left( \frac{f_i}{g} \right)^{N_{ie}-1} \frac{\exp\left(-f_i/g_i M\right)}{\Gamma(N_{ie}) M^{N_{ie}}} \]  (4.10)

One can then incorporate the readout noise via a convolution as was done in Eq. 4.8. In practice however the EM register makes this unnecessary as the readout noise is effectively less than 1 \( e^- \) when compared to the pre-amplified signal. Thus the readout noise will be omitted from the EMCCD noise model in this work as it adds significant complexity to the problem without a corresponding increase in accuracy. Even with omitting the readout noise Eq. 4.10 poses some computational headaches, and can be rather unwieldy to work with. A closed form approximation to the Poisson-gamma distribution has been derived in [123, 128], and was shown to be sufficiently accurate. It is simply stated here as,
\[ p(f | O_0) = \sqrt{\frac{g(qG_i(0)) + c_{sp}}{fM}} \exp\left(-\frac{f}{M} - g(qG_i(0)) + c_{sp}\right) I_1 \left(2\sqrt{\frac{fg(qG_i(0)) + c_{sp}}{M}}\right) \]  (4.11)

where \( I_1 \) is a modified Bessel function of the first kind. The Poisson-gamma distribution is characterized by its skew towards high electron counts commonly referred to as the spurious charge tail.

Equations 4.9 and 4.11 will be used to formulate the Bayesian deconvolution problems presented in the following chapter. While not the most exact formulations they represent a reasonable balance between quantitative accuracy and ease of computation. Proper consideration of the imaging and noise model is essential to implementation of photometrically precise post-processing schemes which are becomingly increasingly important in biological imaging studies.
Thus far, towards the goal of creating an optimal quantitative EDF hybrid imaging for optical microscopy, the focus has been on the modeling, design, and optimization of the imaging system. However this only represents half of the problem, and thus it is also essential to address the computational aspects of the system. As described in Chapter 1, once an extended depth field image has been acquired it is often necessary and/or beneficial to apply a deconvolution filter in order to restore contrast and remove noise. There are several approaches to accomplishing this goal, but perhaps none as prominent in the literature as maximum likelihood estimation and the method of least squares. The most common implementation of this method in microscopy has been the locally varying linear Wiener filter [126, 129]. While it is generally easy to implement and computationally very efficient, the Wiener filter suffers from fundamental trade-offs between resolution and noise amplification in the deconvolved images. More recently $L^1$ regularized deconvolution models have grown in popularity as a method for overcoming this trade-off. Total variation (TV) regularization [130] has been demonstrated as an easily implementable model which shows superior results to the Wiener filter when certain sparsity assumptions are valid. Specifically, this approach has been shown to deconvolve images and suppress noise while maintaining sharp discontinuities (i.e. edge features) and other high spatial frequency image features. It and other $L^1$ regularized models are considered by many to be the heirs apparent to the linear techniques, and have been successfully demonstrated in a number of imaging applications [131, 132, 133, 134, 135].

While still an active area of research, the mathematical foundations of $L^1$ regularized
Deconvolution algorithms are now well established [135, 136, 137, 138, 139, 140, 141, 142, 143]. Thus in keeping with the applications oriented focus of this thesis, a primary goal of this chapter is to evaluate these algorithms with respect to ease of use, speed, and accuracy. This is done in consideration of the end users of EDF systems who, generally speaking, are more concerned with the biological question than the mathematical framework. Ideally a deconvolution software package would operate successfully requiring little to no input from the end user. Another primary concern is the speed of implementation with the ultimate goal of real time processing. Despite the continued validity of Moore’s Law, only the most efficient algorithms for solving the TV deconvolution model have the potential to reach real time implementation. This is especially true as focal plane arrays increase in size (up to 5.5 megapixels in some cases). In this chapter two optimization algorithms are investigated for solving the TV regularized deconvolution problem: the split-Bregman algorithm (SBA) [140, 141, 142] and the fast iterative shrinkage-thresholding algorithm (FISTA) combined with expectation maximization TV (EM-TV) [135, 136, 137, 144].

In addition to a comparison of deconvolution models and minimization algorithms, the EDF pupil functions described in Chapter 3 are compared in simulation to evaluate their performance after deconvolution. The investigation to follow utilizes the EDF PSF analysis presented in Chapter 3 along with microscopy specific noise models presented in Chapter 4 to form the first quantitative comparison of EDF hybrid imaging in high NA microscopy systems. While the use of hybrid imaging to extend the DOF is prevalent in the literature, its application to microscopy has been somewhat limited. Additionally the corresponding analysis of these investigations has always been somewhat narrow in scope, usually only focusing on one aspect of the system in sufficient detail. Arnison provided a theoretical framework for modeling wavefront coded imaging systems under high NA vectorial conditions, and showed some experimental results with the cubic phase mask [35]. However, the work featured little to no discussion of deconvolution or noise models. Chu developed a scalar high NA model for EDF imaging using logarithmic aspheric lenses, and proposed using a total variation regularized deconvolution model [88]. However her analysis assumed a simplified uniform Gaussian noise model, and only showed results for simple, block like objects which were
not representative of typical scenes. Abrahamsson utilized a high NA vectorial imaging model along with appropriate noise and deconvolution models, however the work did not include any quantitative analysis [89]. The vast remainder of the EDF wavefront coding literature focuses on low NA systems and overly simplified noise models [60, 145, 146, 147, 148, 149, 150].

To summarize, the goals of this chapter are three-fold: (1) formulate deconvolution models for EDF imaging in the context of high NA optical microscopy, (2) analyze the various advantages and disadvantages of relevant minimization algorithms from the perspective of the end user, and (3) provide a comparative assessment of EDF images processed via total variation regularized deconvolution for the various pupil masks presented in Chapter 3.

5.1 Isotropic TV Regularized Deconvolution Model for CCD and sCMOS Cameras

A Bayesian framework is used to formulate the TV regularized deconvolution model tailored for EDF images captured with either charged coupled device (CCD) or scientific complementary metal-oxide semiconductor (sCMOS) cameras [130, 151, 152]. The complete imaging system at a given pixel, $i$, can be modeled as the following linear, shift-invariant system

$$ f_i = \mathcal{P}\left(g_i \cdot (q \cdot G_i(O_0) + c_{sp,i})\right) + \mathcal{N}\left(0, (g_i \sigma_i)^2\right), \quad (5.1) $$

where $f$ represents the acquired noisy EDF intensity image in units of image counts, $\mathcal{P}(\cdot)$ denotes a Poisson random variable, and $\mathcal{N}(\mu, \sigma^2)$ denotes a normal distribution of mean, $\mu$, and variance, $\sigma^2$. It is assumed that $f$ is bounded and positive. Recalling the noise model developed in Chapter 4, $g$ is the analog to digital gain factor, $q$ is the quantum efficiency of the sensor, $c_{sp}$ is the spurious charge, and $G_i(O_0)$ is the function that samples the original object, $O_0$, convolved with the intensity PSF, $h$.

The likelihood function described by Eq. 5.1 is written as the discrete convolution of a Poisson variable with a Gaussian variable. However it will instead be approximated as a purely Poisson random variable in order to simplify computation as detailed in the previous chapter. By
recalling that a Gaussian random variable with mean \((g\sigma)^2\) and variance \((g\sigma)^2\) asymptotically approaches a Poisson random variable of mean \((g\sigma)^2\) the imaging model can be written as,

\[ f_i \approx P(y_i), \quad (5.2) \]

where,

\[ y_i = g_i \cdot (q \cdot G_i(O_0) + c_{sp,i}) + (g_i\sigma_i)^2. \quad (5.3) \]

This approximation was shown to be valid and effective by Huang et al. in localization microscopy applications provided sufficiently high signal [127]. These criteria are met for the simulated and experimental results described below. Note that in the case of CCD cameras \(g_i\) and \(\sigma_i^2\) are constants instead of pixel dependent quantities.

The deconvolution problem can then be formulated from the Bayesian perspective of maximum a-posteriori estimation [153]. The problem can be stated as: find the estimate, \(O\), of the object, \(O_0\), given an observed noisy image \(f\), that maximizes the posterior probability distribution. Bayes’ theorem can be used to write an expression for this probability,

\[ P(O|f) = \frac{P(f|O)P(O)}{P(f)}. \quad (5.4) \]

\(P(f|O)\), is the likelihood function and describes the noise statistics of the observed image. \(P(O)\), is the Bayesian prior, and serves to regularize the deconvolution. \(P(f)\), is the evidence and is a constant with respect to the variable being estimated. As such it will not factor into the deconvolution.

It is assumed that the noise is independent of the pixels. Thus using the approximation of Eq. 5.2, the likelihood function for the entire image can be written as the product of the likelihood function of a Poisson variable for each pixel,

\[ P(f|O) = \prod_i \frac{e^{-y_i}y_i^{f_i}}{f_i!}. \quad (5.5) \]

Note as the inverse problem is now under consideration, \(O_0\) is replaced by \(O\) in the variable \(y_i\) for Eq. 5.5 and all following instances.
Now considering the Bayesian prior, the isotropic total variation term [130, 151] is introduced to regularize the imaging model. It is defined as

\[ P(O) = \frac{1}{K} \exp \left( -\beta \sum_i \| \nabla O_i \|_2 \right), \]  

(5.6)

where for an image comprised of \( M \times N \) pixels, \( \nabla \in \mathbb{R}^{M,N} \) represents the discrete gradient operator, \( \| \cdot \|_2 \) represents the \( L^2 \) norm, \( K \) is a normalization constant, and \( \beta \) is a positive constant regularization parameter. The total variation term imposes a smoothness constraint that is able to preserve edge features in the image. This regularizer can lead to a significant improvement over the classic trade-off between resolution and processing artifacts commonly seen in linear filters when the spatial intensity derivative of the image is sparse.

The goal of the deconvolution is to find the estimated object, \( O \), that maximizes Eq. 5.4, or equivalently minimizes its negative logarithm. Ignoring constants with respect to optimization variable \( O \), Eqs. (5.5) and (5.6) are substituted into \(-\log(P(f|O)P(O))\) to yield the following minimization problem,

\[
\min_O \sum_i (y_i - f_i \log y_i) + \beta \sum_i \| \nabla O_i \|_2, 
\]  

(5.7)

where the \( \beta \) regularization parameter serves to balance the weight between the data fidelity and TV denoising terms. Equation (5.7) will be referred to as the TV-Poisson model henceforth.

### 5.2 Minimization Algorithms

The TV-Poisson model, Eq. 5.7, is an unconstrained, \( L^1 \) regularized, convex optimization problem. Common techniques to solve convex problems, such as interior point methods [153], are very computationally expensive and are often intractable for large scale image processing applications. A number of algorithms [131, 132, 134, 138] have been proposed to solve the \( L^1 \) problem more efficiently. Two such algorithms are investigated here: the fast iterative shrinkage-thresholding algorithm coupled with expectation maximization total variation (FISTA/EM-TV) and the split-Bregman algorithm (SBA). While these two methods are very similar, as both can be classified as proximal splitting algorithms, there are some differences that affect their ease of use and speed of
convergence. FISTA solves the TV model without modification and thus only requires the user to specify one regularization parameter. SBA solves a slightly modified version of the TV model and represents one of the fastest methods for problems of this type. However, the modified problem solved by SBA requires the user to select and balance multiple regularization parameters to solve the problem most efficiently which can complicate its use. Both algorithms will reach the same answer given enough iterations, however the two implementations give the option of ease of use or speed of solution when solving the deconvolution problem. A brief description and outline of both algorithms follows.

5.2.1 FISTA/EM-TV

FISTA [139, 144] is an improvement to the original iterative shrinkage-thresholding algorithm (ISTA) [154, 155]. ISTA is an iterative fixed point forward backward splitting algorithm that consists of two steps per iteration. The roots of this algorithm can be traced back to the proximity operator introduced by Moreau [156]. An overview of proximal splitting methods can be found in [143]. Each iteration first takes a step towards minimizing the differentiable data fidelity term and then takes a step towards minimizing the TV regularization term, which is non-differentiable. The data fidelity term is minimized via a standard gradient descent step. The TV regularization term is minimized via its proximity operator, prox_t(·), which is calculated by the EM-TV algorithm [135, 144]. The EM-TV algorithm used is analogous to Chambolle’s algorithm [136, 137] for the case of Poisson noise. The algorithm approximates the proximity operator to within some tolerance ε_2. As discussed in [144], the inner EM-TV loop should be allowed to iterate a minimum of 10 times to ensure convergence of FISTA/EM-TV algorithm as a whole. Since the model, Eq. 5.7, is convex there are no local minimizers, only global ones. The FISTA/EM-TV algorithm is thus guaranteed to generate a sequence J(O^k) that will converge towards the optimal value of J(O)
until some termination criteria has been met, where,

\[ J(\mathcal{O}) = J_{\text{err}}(\mathcal{O}) + \beta J_{\text{reg}}(\mathcal{O}) \]  
(5.8)

\[ J_{\text{err}}(\mathcal{O}) = \sum_i (y_i - f_i \log y_i) \]  
(5.9)

\[ J_{\text{reg}}(\mathcal{O}) = \sum_i \|\nabla \mathcal{O}_i\|_2. \]  
(5.10)

The ISTA algorithm can be seen in Table 5.1 where the Lipschitz constant in this case is 2 times the smallest eigenvalue of the blurring operator, \( h \), in matrix form.

<table>
<thead>
<tr>
<th>Table 5.1: ISTA for the TV-Poisson model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initialize ( L ):= The smallest Lipschitz constant of ( \nabla J_{\text{err}} ), ( k = 1 ), and ( \mathcal{O}^0 = f )</td>
</tr>
<tr>
<td>2. Iterate ( k ), While ( |\mathcal{O}^{k+1} - \mathcal{O}^k|_2/|\mathcal{O}^k|_2 &gt; \epsilon_1 )</td>
</tr>
<tr>
<td>Find ( \mathcal{O}^k ) via EM-TV s.t. ( |\mathcal{O}^k - \text{prox}<em>{\beta/L}(\mathcal{O}^{k-1} - \frac{1}{L} \nabla J</em>{\text{err}}(\mathcal{O}^{k-1}))| &lt; \epsilon_2 )</td>
</tr>
<tr>
<td>3. End</td>
</tr>
</tbody>
</table>

The modification which converts ISTA to FISTA [139] utilizes Nesterov’s accelerated gradient descent [157] to add a momentum term, \( t_k \), and achieve an optimal rate of convergence. The modified FISTA algorithm is shown in Table 5.2. The EM-TV algorithm used to perform the proximity operation also makes use the momentum term [135, 144]. FISTA outperformed both ISTA and the two-step iterative shrinkage-thresholding (TwIST) algorithm [158], consistently achieving a lower value for the functional being minimized in fewer iterations.

<table>
<thead>
<tr>
<th>Table 5.2: FISTA for the TV-Poisson model</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initialize ( L ):= The smallest Lipschitz constant of ( \nabla J_{\text{err}} ), ( k = 1 ), and ( \theta^1 = f )</td>
</tr>
<tr>
<td>2. Iterate ( k ), While ( |\mathcal{O}^{k+1} - \mathcal{O}^k|_2/|\mathcal{O}^k|_2 &gt; \epsilon_1 )</td>
</tr>
<tr>
<td>(a) Find ( \mathcal{O}^k ) via EM-TV s.t. ( |\mathcal{O}^k - \text{prox}<em>{\beta/L}(\theta^{k-1} - \frac{1}{L} \nabla J</em>{\text{err}}(\theta^{k-1}))| &lt; \epsilon_2 )</td>
</tr>
<tr>
<td>(b) ( t_{k+1} = \frac{1+\sqrt{1+4t_k^2}}{2} )</td>
</tr>
<tr>
<td>(c) ( \theta_{k+1} = \mathcal{O}<em>k + \left(\frac{t_k-1}{t</em>{k+1}}\right)(\mathcal{O}<em>k - \mathcal{O}</em>{k-1}) )</td>
</tr>
<tr>
<td>3. End</td>
</tr>
</tbody>
</table>

It is of course necessary to determine the optimal value of \( \beta \) for a given image and/or
application. This can be determined from a simple parameter search using a representative image, and adjusted as necessary for particular noise conditions. Having only one regularization parameter makes this a relatively reasonable task that does not significantly impede generating a suitable EDF image for further data analysis. Given its relative simplicity and speed, FISTA represents an easily implementable minimizer of the TV regularized deconvolution problem.

5.2.2 Split-Bregman Algorithm

The split-Bregman algorithm similarly makes use of proximity operators to solve $L^1$ regularized problems, however it modifies the problems in order to achieve a faster rate of convergence. The algorithm is closely related to the alternating direction method of multipliers (ADMM) and Douglas-Rachford splitting, as noted in [159] and the references therein. Specifically it was noted that when the constraints on SBA are linear it is equivalent to ADMM and thus can make use of existing convergence proofs to justify use of the algorithm. The SBA operates by introducing two slack variables, $\nabla O \rightarrow d$ and $y \rightarrow z$, into Eq. 5.7. The $d$ variable is an $MN \times 2$ vector consisting of the differential operator applied in both transverse spatial dimensions to yield $d = (d_x, d_y)$. The slack variables are kept close to their intended values by the addition of two quadratic penalty functions into the optimization problem. The deconvolution can then be rewritten as the following unconstrained minimization problem,

$$\min_{O,d,z} \sum (z - f \log z) + \beta \sum \| (d_x, d_y) \|_2 + \frac{1}{2\gamma_1} \| d - \nabla O - b_1 \|_2^2 + \frac{1}{2\gamma_2} \| z - y - b_2 \|_2^2. \quad (5.11)$$

Two positive constant regularization parameters, $\gamma_1$ and $\gamma_2$, are introduced to balance the weight of the penalty terms. The update variables, $b_1^k$ and $b_2^k$, iteratively evolve the estimates for $(O,d,z)$. The $b_1^k$ update variable, like $d$, consists of $x$ and $y$ components such that $b_1^k = (b_1^k(x), b_1^k(y))$. Problems of the form of Eq. 5.11 are of a penalty-type, and the classical penalty methods solve this for a sequence of $\gamma_1$ and $\gamma_2$ which approach zero, and thus more and more enforce the constraints. This approach has generally fallen out of favor since the problem becomes more ill-conditioned as $\gamma_1$ and $\gamma_2$ become small. The split-Bregman method instead allows the $\gamma$ parameters to be fixed
and instead varies $b_1^k$ and $b_2^k$ which keeps the problem well conditioned. While Eq. 5.11 appears more complex than the original problem, the introduction of the two new penalty functions allows the optimization to be broken down into three distinct sub-problems and two update equations which can be solved by the iterative algorithm shown in Table 5.3. The $O$ and $z$ sub-problems can be easily solved with standard gradient methods [141, 142]. The $d$ sub-problem, is solved by the proximity operator, which in this case is a shrinkage/soft-thresholding operation [140, 141, 142].

Table 5.3: Split-Bregman algorithm for the TV-Poisson model

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initialize</td>
<td>$k = 0$, $O^0 = f$, $d^0 = \nabla f$, $z^0 = f$, and $b_1^0 = b_2^0 = 0$</td>
</tr>
</tbody>
</table>
| 2. Iterate $k$, While | $\|O^{k+1} - O^k\|_2/\|O^{k+1}\|_2 > \epsilon_1$
  |   | (a) $O^{k+1} = \arg \min_{O} \frac{1}{2\gamma_1} \| \nabla O - d^k + b_1^k \|_2^2 + \frac{1}{2\gamma_2} \| y - z^k + b_2^k \|_2^2$
  |   | (b) $d^{k+1} = \arg \min_d \sum d_{xz} + \frac{1}{2\gamma_1} \| d - \nabla O^{k+1} - b_1^k \|_2^2$
  |   | (c) $z^{k+1} = \arg \min_z \langle 1, z - f \log z \rangle + \frac{1}{2\gamma_2} \| z - y^{k+1} - b_2^k \|_2^2$
  |   | (d) $b_1^{k+1} = b_1^k + (\nabla O^{k+1} - d^{k+1})$
  |   | (e) $b_2^{k+1} = b_2^k + (y^{k+1} - z^{k+1})$
| 3. End | |

The values for the regularization parameters are again determined empirically. As with the FISTA/EM-TV algorithm the primary parameter $\beta$ must be small enough to ensure sufficient resolution and quantitative accuracy, but not so small as to render the total variation smoothing term negligible. The $\gamma_1$ and $\gamma_2$ parameters do not influence the end solution, but rather control the rate of convergence of the algorithm. They are selected to yield accurate solutions within a reasonable number of iterations. In practice it can be difficult to find a proper balance of the regularization parameters so as to yield a reasonable solution with a sufficiently fast speed of convergence. One can again implement a parameter search, however the addition of the two $\gamma$ parameters significantly increases the search space. Thus it can take considerably longer to find the proper balance of the parameters somewhat negating the speed benefits of SBA. Once the optimal values for the parameters have been determined for a given set of imaging conditions, the split-Bregman algorithm represents one of the fastest, if not the fastest, available solution for the
5.2.3 Quantitative Comparison of Algorithms

A series of numerical experiments were performed in order to compare the relative speed differences and performance between FISTA and SBA. A Bessel beam PSF, Fig. 5.1, was convolved with the camera man image, Fig. 5.2(a), to simulate an EDF image, Fig. 5.2(b). The images are 256×256 with a maximum intensity of 5000 photons. The resulting images from FISTA and SBA can be seen in Figs. 5.2(c) and 5.2(d) respectively.

The images in Figs. 5.2(c) and 5.2(d) were generated by first conducting a search over the $\beta$ parameter to find its optimal value with respect to the power SNR (PSNR) and error function. The PSNR is defined as,

$$\text{PSNR} = 20 \log_{10} \frac{\|O_0 - \hat{O}\|_2^2}{\|O_0 - \hat{O}\|_2^2},$$

(5.12)

where $O_0$ is the original object, $\hat{O}$ is the mean intensity of $O_0$, and $\hat{O}$ is the estimated object from the respective deconvolution algorithm. The boundary condition in the total variation calculation is slightly different between the two algorithms. Thus, for the following comparison, the PSNR
Figure 5.2: Equal PSNR cameraman images for FISTA and the split-Bregman algorithm. The intensity range spans from 0 - 5000 photons for all images.
is taken only over the central $64 \times 64$ region to avoid skewing the results. In the case of SBA the $\gamma_1$ and $\gamma_2$ parameters were fixed to 800000 and 1000 respectively as these values were empirically determined to provide a reasonable rate of convergence. For the purposes of this comparison both the quantum efficiency, $q$, and the A/D gain, $g$, are assumed to be unity, and the spurious charged is assumed to be negligible, $c_{sp} = 0$. The standard deviation of the readout noise is assumed to be $\sigma = 1 \text{e}^-$. Figures 5.3(a) and 5.3(b) show the PSNR and error function plotted against $\beta$. The dashed and solid lines show the results after 100 iterations which represents a termination limit that would be typically used in an experimental setting where results are necessary in (near) real time. FISTA has a maximum PSNR of 27.46 dB and SBA has a maximum PSNR of 27.23 dB. FISTA appears to consistently out perform SBA in terms of PSNR, although this result is tied to the 100 iteration termination criterion. For a fair comparison, the x and o markers in Fig. 5.3 show the results after $1 \times 10^6$ iterations to demonstrate that the two algorithms eventually converge to the same answer given enough iterations. Both algorithms show an increase in the error function value with larger values of $\beta$, although this increase is a nearly negligible percentage of the total error functional value. Thus for the cameraman image under the specified noise conditions the optimal value for the regularization parameter was found to be $\beta = 1.10 \times 10^{-3}$.

Next the convergence speed of the two algorithms was analyzed in both computation time and number of iterations with the $\beta$ parameter was set to the optimal value. The termination tolerance set to $\epsilon_1 = 1.00 \times 10^{-3}$ and the maximum number of iterations was set to 100. The algorithms are compared on a Lenovo Thinkpad T420 laptop computer with an Intel i5-2410M (2.30 GHz) processor and 16 GB of RAM using MATLAB. The PSNR and objective function are plotted against the number of iterations in Fig. 5.4. FISTA terminates after 63 iterations running for 6.011 seconds. The split-Bregman algorithm terminates after 17 iterations taking only 0.288 seconds. Thus the time per iteration is an order of magnitude faster for SBA (0.0169 s/iteration) than for FISTA (0.0954 s/iteration) with SBA requiring $3.7 \times$ fewer iterations. Keeping in mind these computation times are for a laptop computer on MATLAB, near real-time processing could be achieved with faster hardware and working in a lower level language such as C/C++.
Figure 5.3: PSNR vs. $\beta$ (a) and the error function vs. $\beta$ (b) for FISTA and SBA. Results after 100 iterations show difference between algorithms under typical use conditions. Results for 1E6 iterations show convergence of both algorithms to the same solution. The additional regularization parameters for SBA were $\gamma_1 = 800000$ and $\gamma_2 = 1000$.

Figure 5.4: PSNR (a) and objective function (b) comparison for FISTA/EM-TV and the split-Bregman algorithm.
The discrepancy in speed between the two algorithms becomes more significant as the image size increases.

As expected, with proper choice of regularization parameters SBA converges significantly faster than FISTA. However, under what would be considered typical termination criteria for experimental imaging, FISTA achieves a slightly higher PSNR. It should be noted that the performance of SBA, in terms of iterations necessary for convergence, could potentially be improved with additional tuning of the $\gamma$ parameters. Optimal tuning of these parameters can be a time consuming process in and of itself, and needs to be redone for varying levels of noise. Thus, FISTA represents a somewhat simpler solution if the end user desires minimal interaction with the algorithm. Whereas SBA is the more desirable solution if a user requires the utmost speed and is willing to take the time to appropriately set the regularization parameters.

### 5.3 Comparison of EDF Pupil Masks Post-Processing

The EDF pupil masks described in Chapter 3 are compared again to study their performance after deconvolution. The $512 \times 512$ phantom shown in Fig. 5.5(a) is the 2D projection of a 3D object that is taken as the ground truth in the following comparisons. The phantom consists of a variety of structures intended to mimic a biological specimen. The central circle, and features contained within, resides at the plane of best focus. The wedges, which shall be referred to by their respective cardinal direction (e.g. N, SSW, NNE, etc.), extend in depth from -4 $\mu$m to 4 $\mu$m. The SSW wedge lies at -1 $\mu$m, and moving clockwise the following wedges decrease in depth in 1 $\mu$m increments until the N wedge. Similarly the S wedge lies at 1 $\mu$m, and moving counter clockwise the following wedges increase in height in 1 $\mu$m increments until the NNE wedge. The phantom imaged by a 1.4 NA objective is shown in Fig. 5.5(b). The limited DOF significantly blurs features for all regions outside the plane of best focus.

The phantom was then convolved with the the cubic, axicon, IPP15, and IPP7-BPM02 PSFs from the appropriate depth planes to simulate EDF images. The peak intensity of the ground truth object was assigned values of 1000, 2000, 5000, and 10,000 photons. The EDF images for
Figure 5.5: (a) Two-dimensional projection of the ground truth 3D EDF phantom, (b) EDF phantom imaged by 1.4 NA clear aperture objective.
the peak ground truth intensity of 5,000 photons are shown in Fig. 5.6. The images show a significantly longer DOF with all of the phantom features now present in a single frame. The circularly symmetric pupil masks show an image that is roughly similar in shape and structure to the ground truth object just with some blurring. Deconvolution will enhance the resolution and quality of the EDF image, but is not strictly necessary to interpret the intermediate image for the axicon, IPP, or IPP-BPM pupil masks. The intermediate image generated from the cubic phase mask is extremely blurred however, and is of minimal use without post-processing. Additionally, the cubic PSF is spread out over a larger area decreasing the SNR of the intermediate image with respect to the circularly symmetric pupil masks.

The resulting intensities and PSNR of the various pre- and post-processed EDF images are shown in Table 5.4. The preprocessed quantities are denoted by the subscript, i, for initial, and the post-processed quantities are denoted by the subscript, f, for final. The images are processed using the split-Bregman algorithm. For each value of the ground truth peak intensity, the optimal $\beta$ was empirically determined through a parameter search to select the value that yielded highest SNR without over smoothing the image. The additional regularization parameters were set to $\gamma_1 = 500,000$ and $\gamma_2 = 1000$, and the termination tolerance was set to $\epsilon_1 = 1E-3$. The IPP07-BPM02 pupil mask shows the highest value and best relative improvement for the PSNR up to a ground truth peak intensity of 5000 photons. At the highest signal images, where the ground truth peak intensity was 10,000 photons, the axicon shows a slightly better performance than the IPP07-BPM02 pupil mask. Despite its better performance at high intensity levels, it is important to note that the axicon PSF quickly becomes ring shaped on one side of focus. Thus for very thick objects, EDF images generated with an axicon would contain ring shaped artifacts. Conversely, the partitioned PSFs blur uniformly outside the EDF design region and thus would not generate additional artifacts for this samples.

In looking at the post-processed images, one can build a qualitative picture to go along with the quantitative analysis. Figures 5.7 - 5.10 compare the deconvolved phantoms for all 4 masks for ground truth peak intensities of 1000, 2000, 5000, and 10,000 photons. In general, the circularly
Figure 5.6: EDF phantoms generated from ground truth image with peak intensity of 5000 photons for the (a) axicon, (b) cubic, (c) IPP15, and (d) IPP07-BPM02 phase masks. All images are plotted from 0-950 image counts.
Table 5.4: Improvement in PSNR after deconvolution over various ground truth signal intensities.

<table>
<thead>
<tr>
<th>Ground Truth Peak Intensity $N_p = 1000, \beta = 0.0005$</th>
<th>Phase Mask</th>
<th>$N_p$ EDF Peak</th>
<th>PSNR$_i$</th>
<th>PSNR$_f$</th>
<th>$\Delta$PSNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubic</td>
<td>157</td>
<td>15.50</td>
<td>16.02</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>Axicon</td>
<td>174</td>
<td>15.71</td>
<td>18.74</td>
<td>3.03</td>
<td></td>
</tr>
<tr>
<td>IPP15</td>
<td>200</td>
<td>15.85</td>
<td>19.17</td>
<td>3.32</td>
<td></td>
</tr>
<tr>
<td>IPP07-BPM02</td>
<td>196</td>
<td>15.82</td>
<td>19.22</td>
<td>3.40</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ground Truth Peak Intensity $N_p = 2000, \beta = 0.0005$</th>
<th>Phase Mask</th>
<th>$N_p$ EDF Peak</th>
<th>PSNR$_i$</th>
<th>PSNR$_f$</th>
<th>$\Delta$PSNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubic</td>
<td>299</td>
<td>15.51</td>
<td>16.29</td>
<td>0.78</td>
<td></td>
</tr>
<tr>
<td>Axicon</td>
<td>319</td>
<td>15.71</td>
<td>19.55</td>
<td>3.84</td>
<td></td>
</tr>
<tr>
<td>IPP15</td>
<td>378</td>
<td>15.85</td>
<td>19.54</td>
<td>3.69</td>
<td></td>
</tr>
<tr>
<td>IPP07-BPM02</td>
<td>364</td>
<td>15.82</td>
<td>19.77</td>
<td>3.95</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ground Truth Peak Intensity $N_p = 5000, \beta = 0.0001$</th>
<th>Phase Mask</th>
<th>$N_p$ EDF Peak</th>
<th>PSNR$_i$</th>
<th>PSNR$_f$</th>
<th>$\Delta$PSNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubic</td>
<td>709</td>
<td>15.51</td>
<td>16.48</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>Axicon</td>
<td>774</td>
<td>15.71</td>
<td>20.82</td>
<td>5.11</td>
<td></td>
</tr>
<tr>
<td>IPP15</td>
<td>949</td>
<td>15.85</td>
<td>21.80</td>
<td>4.95</td>
<td></td>
</tr>
<tr>
<td>IPP07-BPM02</td>
<td>872</td>
<td>15.83</td>
<td>21.97</td>
<td>5.14</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ground Truth Peak Intensity $N_p = 10,000, \beta = 0.00005$</th>
<th>Phase Mask</th>
<th>$N_p$ EDF Peak</th>
<th>PSNR$_i$</th>
<th>PSNR$_f$</th>
<th>$\Delta$PSNR (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cubic</td>
<td>1410</td>
<td>15.51</td>
<td>16.62</td>
<td>1.11</td>
<td></td>
</tr>
<tr>
<td>Axicon</td>
<td>1531</td>
<td>15.71</td>
<td>21.97</td>
<td>6.26</td>
<td></td>
</tr>
<tr>
<td>IPP15</td>
<td>1843</td>
<td>15.86</td>
<td>21.07</td>
<td>5.21</td>
<td></td>
</tr>
<tr>
<td>IPP07-BPM02</td>
<td>1757</td>
<td>15.83</td>
<td>21.62</td>
<td>5.79</td>
<td></td>
</tr>
</tbody>
</table>
symmetric pupil masks have minimal artifacts when compared to the cubic phase mask. Several ‘ghost artifacts’ can be seen in the deconvolved cubic images regardless of the signal strength. This is the direct result of the asymmetric pupil mask which leads to an accumulation of phase. It is also clear in the deconvolved cubic images that the position of wedges in the phantom far from the traditional focal plane have translated from their true positions. This again can be attributed to the asymmetric nature of the cubic phase mask. Visual inspection of the deconvolved images from the circularly symmetric phase masks supports the quantitative analysis of Table 5.4.

5.4 TV Poisson-Gamma Model for EMCCD Cameras

The deconvolution problem for EMCCD cameras utilizes the Poisson-Gamma noise model described in Chapter 4. Given the advent of comparably inexpensive sCMOS cameras the EMCCD is beginning to fall out of favor for many bio-imaging applications, however they still currently offer the best noise performance for ultra low signal imaging. The imaging model is described as the following Poisson-Gamma process, $\mathcal{P}_G$,

$$f_i \approx \mathcal{P}_G(y_i; M),$$

(5.13)

where,

$$y_i = g_i \cdot (q \cdot G_i(O_0) + c_{sp,i}).$$

(5.14)

Note the omission of the readout noise term in this definition of $y_i$ which is specific to the Poisson-Gamma model, as opposed to Eq. 5.3. This is due to the electron multiplication register which renders the readout noise negligible.

Following the same general framework described for the TV-Poisson model, the TV-Poisson-Gamma model can be written as the negative log likelihood,

$$\min_{\mathcal{O}} \sum_i \left( \frac{f_i}{M} + y_i - \frac{1}{2} \log \frac{y_i}{f_i M} - \log I_1 \left( 2 \sqrt{\frac{f_i y_i}{M}} \right) \right) + \beta \sum_i \| \nabla \mathcal{O}_i \|_2.$$  

(5.15)

As with the case of Poisson noise the above problem is convex. The split-Bregman algorithm is used to solve the minimization as it is easily adapted to the Poisson-Gamma noise model. The
Figure 5.7: EDF Phantoms after deconvolution from a ground truth object emitting a peak intensity of 1000 photons. All images are plotted from 0-1500 photons.
Figure 5.8: EDF Phantoms after deconvolution from a ground truth object emitting a peak intensity of 2000 photons. All images are plotted from 0-3000 photons.
Figure 5.9: EDF Phantoms after deconvolution from a ground truth object emitting a peak intensity of 5000 photons. All images are plotted from 0-6250 photons.
Figure 5.10: EDF Phantoms after deconvolution from a ground truth object emitting a peak intensity of 10000 photons. All image are plotted from 0-12500 photons.
corresponding modified minimization problem is written as,

\[
\min_{\mathcal{O},d,z} \sum \left( f_M \frac{z}{M} - \frac{1}{2} \log \left( \frac{z}{f_M} \right) - \log I_1 \left( 2 \sqrt{\frac{f z}{M}} \right) \right) + \beta \sum \| (d_x, d_y) \|_2
\]

\[
+ \frac{1}{2\gamma} \| d - \nabla \mathcal{O} - b_1^k \|_2^2 + \frac{1}{2\gamma} \| z - y - b_2^k \|_2^2,
\]

and is solved by the algorithm described in Table 5.5. Step 2(c) of the algorithm is solved using a

Table 5.5: Split-Bregman algorithm for the TV-Poisson-Gamma model

<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Initialize</td>
<td>( k = 0, \mathcal{O}^0 = f, d^0 = \nabla f, z^0 = f, ) and ( b_1^0 = b_2^0 = 0 )</td>
</tr>
</tbody>
</table>
| 2. Iterate \( k \), While | \[ \| \mathcal{O}^{k+1} - \mathcal{O}^k \|_2 / \| \mathcal{O}^{k+1} \|_2 > \epsilon_1 \]
| \( a \) | \( \mathcal{O}^{k+1} = \arg \min_{\mathcal{O}} \frac{1}{2\gamma} \| \mathcal{O} - d^k + b_1^k \|_2^2 + \frac{1}{2\gamma} \| y - z^k + b_2^k \|_2^2 \)
| \( b \) | \( d^{k+1} = \arg \min_d \beta \sum \| (d_x, d_y) \|_2 + \frac{1}{2\gamma_1} \| d - \nabla \mathcal{O}^{k+1} - b_1^k \|_2^2 \)
| \( c \) | \( z^{k+1} = \arg \min_z \left( 2 \log \frac{z}{f_M} - \log I_1 \left( 2 \sqrt{\frac{f z}{M}} \right) \right) + \frac{1}{2\gamma_2} \| z - y^{k+1} - b_2^k \|_2^2 \)
| \( d \) | \( b_1^{k+1} = b_1^k + \left( \nabla \mathcal{O}^{k+1} - d^{k+1} \right) \)
| \( e \) | \( b_2^{k+1} = b_2^k + \left( y^{k+1} - z^{k+1} \right) \)
| 3. End | |

basic version of Newton’s method [160] as the variable of interest is not separable. This leads to a considerably slower implementation of the split-Bregman algorithm when compared to the analytic solution that can be used for this step in the TV Poisson model. A more advanced approach with appropriate line search algorithms [160] can be used to improve convergence in applications where speed is critical.

EMCCD cameras that are approximated by the Poisson-Gamma noise model usually observe weakly emitting fluorescent tags that form point like objects. Thus the phantom for this model similarly consists of point like objects scattered throughout the field of view. Figure 5.11(a) shows the diffraction limited axial projection of 9 point emitters. Three emitters are located at the focal plane, 3 are located 2\( \mu \)m below the focal plane, and 3 are located 4\( \mu \)m below the focal plane. Figure 5.11(b) shows the scene imaged by a 1.4 NA clear aperture objective. Significant blurring is observed for the emitters located below the focal plane. The blurred and noisy EDF image, Fig 5.11(c), was simulated assuming an axicon phase mask where the intensity is now presented in
image counts. The Poisson-Gamma noise was simulated using the Monte Carlo rejection sampling method [161]. The EM gain was set to $M = 200$, the A/D gain is assumed to be unity $g = 1$ and spurious charge is assumed to be $c_{sp} = 0.01$. The regularization parameters were empirically determined to be $\beta = 0.0133$, $\gamma_1 = 12.5$, and $\gamma_2 = 0.010$. The deconvolved image is shown in Fig. 5.11(d). The intensity of the point emitters has been successfully estimated and the noise has been significantly mitigated. This is the first presentation and demonstration of an EMCCD specific deconvolution model to the author’s knowledge.

5.5 Conclusions

In this chapter total variation regularized deconvolution models specifically tailored to the sCMOS and EMCCD cameras commonly used in biological imaging were presented. For the sCMOS model, the FISTA/EM-TV and split-Bregman algorithms were compared along with their respective pros and cons. The four EDF pupil masks presented in Chapter 3 were then quantitatively analyzed comparing the post-processed PSNR of an imaging phantom against one another under varying signal strength conditions. The IPP-BPM phase mask was found to yield the highest quality under low and medium intensity signal conditions. The axicon showed the slightly better performance at the highest signal level. The overall structure of the phantom was recovered well across all of the signal strengths, but the corrupting effects of noise were quite evident at lower intensities. The simulations outlined in this chapter can be used as a predictive tool to determine the suitability of hybrid imaging for EDF microscopy under a prescribed set of imaging and camera conditions.

Future work in the deconvolution EDF images should explore several alternatives of the standard model which have been shown to yield improved results. One such extension is reweighted $L^1$ minimization [162] where the problem is solve multiple times and the regularization parameter is reweighted based on the previous solution. The principle behind this is that a more accurate solution can be found if $\beta$ is allowed to vary on a pixel by pixel basis instead of being a single constant for the entire image. The other benefit to reweighting is that it does not bias the total variation to be smaller, which allows for a better estimate of the original object. Other non-convex
Figure 5.11: Deconvolution of simulated imaging data from an EMCCD camera. (a) The diffraction limited projection shows 9 emitters. The colorbar units are in photons. (b) The clear aperture image shows significant blurring from the limited DOF. (c) The blurred and noisy EDF image shows all of the emitters. Note the colorbar units are now in image counts and significantly higher due to the EM gain. (d) The deconvolved image shows an extended DOF, accurate estimation of the photon level, and minimal noise.
alternatives to the standard model include iterative hard thresholding [163] and $\ell_p$ minimization for $0 < p < 1$ [164]. When starting from a good initialization point (e.g. the solution to the standard convex problem) these methods have been shown to yield superior results especially under low SNR conditions.
Chapter 6

Fabrication, Verification, and Applications

Experimental verification is the final stage in the design and development of the proposed IPP-BPM pupil mask for extended depth of field optical microscopy. This can be broken down into three main categories: fabrication and characterization of the pupil masks, verification of experimental performance, and application to a biological problem of interest.

6.1 Fabrication and Characterization

The IPP-BPM phase mask designed in Chapter 3 consists of features on both the micrometer and nanometer scale as indicated in Fig. 6.1. Fabricating such a mask within a single substrate was not a viable option given current methods. Available processes are generally tailored to one scale or the other. Thus rather than attempting to develop a new process to fabricate a single IPP-BPM mask, it was fabricated as two separate components, and optically combined together. The fabrication of the individual IPP and BPM pupil masks is described below.

![Figure 6.1: Sketch of a possible realization of an IPP-BPM phase mask.](image-url)
6.1.1 Fabrication of the IPP Mask

The IPP mask originally proposed by Abrahamsson [89] was fabricated using optical contacting of 200 μm fused silica circular windows. Maintaining concentricity of each window contributes significantly to cost. Fabricating a mask consisting of only 6 tiers via optical contacting costs upwards of 6000USD. This poses a severe limitation on the feasibility of mass production. Thus alternative techniques were explored to reduce the cost and difficulty of fabrication.

A recent development in micro/nano fabrication is the use of femtosecond laser pulses to modify the material properties of dielectrics [165, 166]. As with other 3D printing techniques there has also been tremendous growth in the availability of turn-key femtosecond laser fabrication systems. FEMTOprint S.A. (Switzerland) provides such systems and was selected to provide the initial prototyping services for the IPP phase mask. Their process creates structures in fused silica with claimed tolerance of ±1 μm on feature sizes in all three spatial dimensions. While this would normally be insufficient for traditional optical elements, it is more than sufficient for the micron level dimensions of the tiered IPP mask. The fused silica substrate allows for further polishing in order to obtain optically flat surfaces (up to a claimed value of λ/50), and an anti-reflection coating could theoretically be applied. Perhaps most importantly the FEMTOprint system is capable of producing masks at a significantly higher volume and lower cost than available from optical contacting. After an initial investment of approximately 3000USD, individual masks can be made for 300-500USD depending on the quantity. This recovers the cost advantages of any hybrid imaging system utilizing IPP pupil masks.

A single 6 tiered mask was utilized for the following proof of concept and verification experiments. Future implementations would theoretically consist of multiple masks mounted on a filter wheel so the user could adjust the DOF to SNR trade-off based on a given experiment. The substrate is a circular fused silica window 4.5 mm diameter and 2 mm in thickness. The radius of each tier was scaled according to the equal area principle described in Chapter 3. The largest tier is 7.2 mm to fit the back aperture diameter of the 63×/1.4 NA Zeiss oil immersion objective.
used. A conceptual sketch of the actual mask geometry is shown in Fig. 6.2 and micrographs of the actual part are shown in Fig. 6.3. The specified tier radius values along with the actual values are shown in Table 6.1. On a qualitative level the micrographs show good concentricity of the tiers. The actual radii differed from the prescribed value by as much as 6 $\mu$m.

The surface of the IPP mask was analyzed using a DEKTAK3030 surface profiler over various height scales. While these line scans do not provide a full interferometric analysis, they can provide an initial estimate of the surface characteristics and insight on the expected performance of the mask. Figure 6.4(a) shows the surface profile from the highest tier of the IPP pupil mask over a vertical range of 200 nm. There is a peak to valley variation of over 100 nm over this distance, and the surface is clearly not smooth. Figure 6.4(b) shows a surface line profile from the third tier over
Table 6.1: Comparison of the specified to the fabricated values for the IPP mask. All values are in mm.

<table>
<thead>
<tr>
<th>Radius ((\Sigma_i))</th>
<th>Prescribed</th>
<th>Actual</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>({w_{\Sigma_1}})</td>
<td>1.470</td>
<td>1.471</td>
<td>0.001</td>
</tr>
<tr>
<td>({w_{\Sigma_2}})</td>
<td>2.078</td>
<td>2.082</td>
<td>0.004</td>
</tr>
<tr>
<td>({w_{\Sigma_3}})</td>
<td>2.546</td>
<td>2.551</td>
<td>0.005</td>
</tr>
<tr>
<td>({w_{\Sigma_4}})</td>
<td>2.939</td>
<td>2.944</td>
<td>0.005</td>
</tr>
<tr>
<td>({w_{\Sigma_5}})</td>
<td>3.286</td>
<td>3.292</td>
<td>0.006</td>
</tr>
<tr>
<td>({w_{\Sigma_6}})</td>
<td>3.600</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

vertical scale of 900 nm. There is a peak to valley variation of over 750 nm. Figure 6.4(c) shows a surface line profile over vertical range of approximately 10 \(\mu\)m. The surface variation relative to the tier height is more readily apparent, as well as the slope of the side wall. Unfortunately, these inconsistencies are enough to significantly degrade the performance of the pupil mask, as will be shown. A similar amount of variation was observed in the other tiers. These measurements suggest that significant transverse blurring will be observed in the experimental IPP PSF.

6.1.2 Fabrication of the Binary Phase Mask

The fabrication of binary phase masks is a straightforward and well documented process. The procedure outlined by Jabbour [80] was followed with slight modifications as described below. The mask was fabricated using the Colorado Nanofabrication Laboratory (CNL) at the University of Colorado at Boulder. The complete process is outlined in Fig. 6.5.

The first step in the fabrication process is to clean the fused silica substrate. This removes any organics that may be present and helps with the adhesion of photoresist. The 10 mm diameter and 3 mm thick substrate is removed from its factory sealed packaging (Edmund Optics), submerged in an acetone bath, and sonicated for 5 minutes. Immediately upon removal from the acetone bath the substrate is submerged in isopropyl alcohol followed by a rinse in a deionized (DI) water bath. To remove any remaining organic compounds the substrate is then submerged in a heated bath of Nanostrip® (Cyantek, sulfuric acid and hydrogen peroxide formulation) for 10-15 minutes, and
Figure 6.4: Surface profiles of the IPP pupil mask at (a) fine, (b) medium, (c) coarse height scales.
Cleaning Solutions: Acetone and Nanostrip

Spin coat - NR71-1500PY for 40 s

Soft bake - 160°C for 1 min

Photoexposure 5 min @ 1.3 mW/cm²

Hard bake - 110°C for 1 min

Development - RD6 10 s

Descum - 40 s O₂ Reactive Ion Etch

Chromium Coating

Acetone Lift Off

CF₄ Reactive Ion Etch (42 nm/min)

Chromium Etch

Figure 6.5: Fabrication process for BPM mask.
then rinsed in a DI water bath for 5 minutes.

The cleaned substrate is now ready to be coated in a layer of photoresist. Negative resist NR71-1500PY (Futurrex) is spun onto the substrate for 40 seconds at 4000 revolutions per minute (rpm). The resulting layer is approximately 1.3 µm thick. After spin coating the substrate is soft baked at 160°C for 1 minute. This is 10°C hotter than prescribed for silica wafers due to the thickness of the substrate. The soft bake improves adhesion, prevents bubbling, and minimizes dark erosion of the photoresist in the following processing steps.

The coated substrate is now ready to be patterned via exposure to ultraviolet light. The transverse profile for the BPM mask is first etched into a chrome lithography mask that was created using a Heidelberg DWL 66FS mask generator. The chrome mask was designed using Clewin4 layout software. A Karl Suss MJB3-HP mask aligner was used to precisely align the chrome mask with the substrate. The sample was then exposed to a 1.3 mW/cm², 365 nm source for approximately 5 minutes to cure the photoresist and establish cross-linking. The substrate is then hard baked at 110°C for 1 minutes to catalytically complete the photo reaction, and reduce mechanical stress on the resist. The substrate is next developed for 10 seconds in RD6 which removes any portions of the resist that were masked during the exposure step to create a negative of the desired transverse profile. To remove any resist that remained after the development process the substrate is then descummed in a reactive ion etcher (March Jupiter III RIE) with an O₂ plasma for 40 seconds.

The next step in the process is to evaporate a chromium coating onto the substrate to serve as a mask during the fused silica etching process. The chrome mask needs to be of sufficient thickness to allow the substrate to reach necessary depth for a π phase shift at the design central wavelength. For the 508 nm emission peak of GFP this equates to a mask thickness of ~550 nm. The substrate is placed in a vacuum chamber that is pumped down to 2×10⁻⁶ Torr. Current is then passed through a chromium rod which evaporates ions on to the substrate at a rate of approximately 1 Å/s. The chrome coated substrate is then sonicated in an acetone bath for 5 min to lift off the remaining photoresist and create the final mask.
The binary phase steps are made by placing the chrome masked substrate in a RIE (Plasmatherm 540). A plasma of 16 standard cubic centimeters per minute (sccm) CF$_4$ and 4 sccm O$_2$ etches the fused silica at a rate of $\sim 41$ nm/min and etches the chrome at a rate of $\sim 3$ nm/min. Assuming central wavelengths of 508 nm and 600 nm the refractive index of fused silica is 1.4619 and 1.4580 respectively. Thus etch depths of 549.9 nm and 655.0 nm achieve the necessary $\pi$ phase shift as calculated via Eq. 3.8. Once the etch is performed the final step in the process is to remove the remnants of the chrome mask using a wet etch process. In practice it was found that a layer of either chromium fluoride or chromium oxide formed during the aforementioned RIE. Thus it was necessary to first submerge the substrate in a heated Nanostrip® bath to remove this layer prior to the chromium etch.

The BPM masks created by this process, Fig. 6.6, were characterized using three different instruments. A DEKTAK3030 was used for measuring the etch depth, an optical microscope was used for measuring the radius of each annular section, and an atomic force microscope (AFM) was used for measuring the sidewall slope. Two BPM masks were fabricated for central wavelengths in the green (508 nm) and the red (600 nm). The DEKTAK used for the depth measurements features Angstrom level precision. The average etch depth was found to be 544.8±5.9 nm for the green mask,
Table 6.2: Comparison of the specified to the fabricated values for the binary phase mask. All values are in $\mu$m.

<table>
<thead>
<tr>
<th>Annulus</th>
<th>Designed Width</th>
<th>Actual Width</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>268</td>
<td>287</td>
<td>19</td>
</tr>
<tr>
<td>2</td>
<td>111</td>
<td>131</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>85</td>
<td>109</td>
<td>24</td>
</tr>
<tr>
<td>4</td>
<td>72</td>
<td>97</td>
<td>25</td>
</tr>
<tr>
<td>5</td>
<td>63</td>
<td>86</td>
<td>23</td>
</tr>
<tr>
<td>6</td>
<td>57</td>
<td>78</td>
<td>21</td>
</tr>
</tbody>
</table>

and 646.1±9.5 nm for the red mask. This equates to a $\pi$ phase shift at a central wavelengths of approximately 503 nm and 592 nm. This is sufficiently close to the design wavelengths so as not to significantly degrade performance.

The width of the binary phase zones determines the flatness of the axial PSF through focus. The phase mask shown in Fig. 6.6 was designed to have a 15% ripple in the axial intensity response over the ‘flat’ region. To achieve this ripple, each annular $\pi$ phase zone was scaled to be $w_{\text{rel}=0.818}$ of the pupil partition tier width. A Zeiss upright microscope was then used to measure the width of each annulus. The Zeiss microscope was calibrated using ruled micrometer slide. These measurements were acquired in bright field using a combination of a 0.08 NA objective for the first annulus, and a 10x/0.25 NA objective for the rest of the annuli. The low power objectives are necessary to achieve a sufficient FOV even though they have reduced resolution. Values are only specified to micrometer accuracy due to both the lower resolution objectives and the feature size limits on the lithography mask. The designed and measured width for each of the 6 etched phase rings is shown in Table 6.2.

As can be seen, there is a consistent difference of 22 $\mu$m on average for all of the annuli. This discrepancy can be attributed to deviations from the ideal in both the printed lithography mask, the photoresist development process, the chrome deposition, the photoresist lift off, and the final plasma etching. While this is a fairly large discrepancy from the design, it should simply yield a larger ripple in the on-axis intensity and possibly produce a larger spot size. Thus the mask
fabricated will be able to serve sufficiently to prove the concept of adding binary phase modulation to incoherent pupil partitioning.

According to the tolerance analysis of binary phase masks presented by Jabbour [80] it was shown that a sidewall slope less than 45° has a minimal effect on performance. Reactive ion etching can theoretically achieve very steep sidewalls in the substrate, however this is dependent on the quality of the sidewalls of the chrome mask as well. A Nanosurf Easyscan2 AFM was used to measure the slope. The amplitude map and depth map are shown in Figs. 6.7(a) and 6.7(b). The line profile for the depth map is shown in Fig. 6.7(c). Analysis of the line profile showed a sidewall slope of approximately 45°. This is relatively shallow sidewall roll off, and right at the edge of the acceptable tolerance range. Thus some additional transverse blurring can be expected.

With proper time, funding, and process tuning the fabrication of the BPM pupil mask can be brought further in-line with the designed specifications for future experiments. However characterization of the etch depth, annulus width, and sidewall slope showed all to be within acceptable values for the proof of concept experiment. Thus while the fabricated BPM mask is far from ideal, it will serve as a useful vehicle to demonstrate the concepts simulated in Chapter 3 of this thesis.

6.2 Experimental Setup and Verification

A Zeiss Axiovert 100M inverted microscope was modified to experimentally evaluate the performance of the composite IPP-BPM pupil mask. A diagram of the experimental setup is shown in Fig. 6.8. Two 4-f relays, utilizing 170 mm achromatic doublets (OptoSigma), were added to the system after the tube lens. The IPP mask was placed at the Fourier plane of the first relay, and the BPM mask was placed at the Fourier plane of the second relay. Both masks were mounted with 5 degrees of freedom (x, y, z, tip, and tilt) so as to allow proper control over the alignment.

PSF measurements were made for the following three cases: a traditional system (i.e. no EDF pupil masks), just the IPP mask, and the full the IPP-BPM system. The measurements were made using a dilution of 100 nm diameter red fluorescent beads excited at 580 nm and emitting
Figure 6.7: AFM profiles taken from a BPM pupil mask early in the tuning of the fabrication process.
Figure 6.8: Experimental setup for EDF fluorescence microscope.
at 605 nm (Life Technologies FluorSpheres). The sensor was an Andor Zyla 5.5 sCMOS camera. The root mean square read noise over the entire detector was $1.2 \times e^{-1}$ and a dark current of $0.001 e^{-1}$. The quantum efficiency was approximately 55% for the green wavelengths and 60% in the red. For each measurement the 63x/1.4 NA objective was scanned through focus in 100 nm increments. The entire system was controlled by the Micro-Manager open source microscopy software package which utilizes ImageJ. Zeiss immersion oil of index $n_{\text{oil}} = 1.518$ was used for all of the following experiments described.

6.2.1 Clear Aperture PSF

The measured axial and transverse intensity profiles for the clear aperture PSF are shown along side the corresponding simulations for a 1.4 NA aberration free objective in Fig. 6.9. As can be clearly seen the axial PSF deviates from the simulations in that it is highly asymmetric through focus as well as many times broader. This indicates the presence of spherical aberration. It is well known that high NA objectives are specifically designed to minimize aberrations only at the focal plane, often times at the expense of significant aberrations elsewhere through focus. Looking at the transverse spot size, it is evident that even at the focal plane the resolution is worse than simulated. As noted in [167], the NA of objectives is often over reported. While interferometric measurements could not be made to explicitly verify the true NA of the objective, the PSF measurements strongly indicate that the actual operating NA of the objective is lower than reported.

In order to determine the best fit parameters, the simulations were run over a range of NA and spherical aberration values. Figure 6.10 compares the experimental data to a simulated PSF from an objective with an NA of 1.15 and 1.53 waves of spherical aberration. As can be clearly seen there is significantly better agreement between the two. The spherical aberration can be attributed to a number of different sources. There is a refractive index mismatch between the immersion oil and the aqueous solution of the fluorescent beads. There is almost certainly some amount of spherical aberration from the optical design of the objective itself. Finally any deviation of the immersion oil index of refraction from the value designed for in the objective would also
induce spherical aberration.

There is still some discrepancy between the simulated and experimental data which indicates the presence of other aberrations within the imaging system. An interferometric analysis is necessary to conclusively determine all of the aberrations present and their sources within the system.

Regardless, the best fit simulation sets the starting point to which the EDF phase masks can be compared. The longer DOF observed in the traditional system should be expected to translate over to the EDF systems as well. Under aberration free imaging at the reported NA (1.4), the IPP phase mask alone was expected to have a DOF of 2.79 µm and the IPP-BPM system was predicted to have a DOF of 5.86 µm. Running the simulations under the new parameters the experimental systems are expected to have a depths of field of 6.11 µm and 12.69 µm respectively.

6.2.2 IPP PSF

The axial and transverse intensity profiles for the PSF measurements of just the IPP mask along with the best fit simulation are shown in Figs. 6.11(a) and 6.11(b). The axial component of the experimental PSF matches very well with simulation. The experimentally measured DOF is 6.12...
Figure 6.10: Best fit simulation to the experimental PSF with NA = 1.15 and 1.53λ of spherical aberration. Transverse (a) and axial (b) line profiles now show much better agreement with experiment. The r-z cross sections of the experimental (c) and simulated (d) PSFs clearly show the effect of spherical aberration. Note the experimental PSF has been shifted from the previous figure.
\[\mu m, \text{ a difference of } 0.01 \mu m \text{ from predicted. The transverse width of the experimental PSF is } 0.572 \mu m, 1.77\times \text{ broader than the predicted value of } 0.322 \mu m. \text{ The broadening is to be expected based on the surface profile measurements previously discussed, and unfortunately is enough to mitigate the EDF benefits of the pupil mask as the transverse PSF is now on par with a clear aperture stopped down to achieve the same DOF. The } x-z \text{ cross section of the measured and simulated PSFs are shown in Figs. 6.11(c) and 6.11(d) respectively. The PSF exhibits a slight bend and asymmetry through focus which can be attributed to the fabrication errors.}

Since a full interferometric analysis of the pupil mask is not currently feasible given budgetary and equipment constraints, the simulation was empirically adjusted to better approximate experimental observations. A combination of corner rounding and random phase variation were added to each tier. The peak to valley phase delay of the random variation was set to 1 wave as estimated from the surface variation line plots shown in Fig. 6.4. The corner rounding, as observed in Fig. 6.4(c), was estimated using high order even polynomial normalized to the outer radius of the given tier. The PSF most consistent with the experimental data can be seen in Fig. 6.12. While there is a still a need for a more quantitatively rigorous analysis, it provides initial insight into the scale and effect of aberrations on the IPP pupil mask. The simulation can be refined by interferometric testing and fitting the resulting interferogram to the annular Zernike polynomials [168].

### 6.2.3 IPP-BPM PSF

The PSF measurements were repeated with both the IPP and the BPM pupil masks in place. Figure 6.13 shows the corresponding axial and transverse line profiles along with the r-z cross sections for both simulated and experimentally measured PSFs. The simulated PSF includes the aberrations previously discussed. The BPM pupil mask was modeled using the experimentally measured values for the phase annuli. The axial intensity shown follows simulation quite well and shows an experimental DOF of 12.76 \(\mu m\) compared to the predicted value of 12.69 \(\mu m\). The r-z cross section shows a combination of fabrication and misalignment error causing the PSF to
Figure 6.11: Comparison of simulated and experimental IPP PSF. The axial line profile (a) matches well, however the experimentally observed transverse profile (b) is significantly broader than predicted. The experimental (c) and simulated (d) r-z cross sections illustrate the loss in transverse resolution from mask fabrication errors.
Figure 6.12: IPP pupil with pseudo-random phase variations to better simulate the experimentally observed PSF. The transverse PSF line profile (a) and r-z cross section (b) now show much better agreement with experiment. Variations on the order of a wave is sufficient to dramatically increase the transverse FWHM of the IPP PSF.
‘wobble’ slightly through focus. The experimentally observed transverse FWHM of 0.994 µm is significantly larger than the predicted value of 0.484 µm. This indicates additional surface level fabrication errors in the BPM pupil mask and/or alignment errors. As with the IPP pupil mask, the resolution is on par with a clear aperture PSF stopped down to achieve the same depth of field.

Despite issues with pupil mask fabrication, the work shown is the first ever experimental demonstration of combined pupil partitioning and spatial wavefront modulation for PSF engineering. This added degree of freedom opens a new space to design pupil functions for extended DOF imaging and beyond. The experimental results can be brought more in-line with theoretical predictions with additional investment and refinement of the fabrication process.

6.3 Biological Applications

Both of the IPP and IPP-BPM configurations of the extended depth of field microscope were evaluated in biological imaging experiments. Due to the discrepancies between the theoretical and experimental PSF measurements, these experiments are analyzed from a qualitative as opposed to quantitative perspective. In the first imaging experiment axons of the Purkinje neuron are observed with the same objective used in the PSF measurements. The cells are labeled with Alexa 488, are excited with blue light and fluoresce in the green. A version of the binary phase mask etched for a central wavelength of 503 nm was used for this experiment. The axons span over 3 µm in depth. Figures 6.14(a) & (b) show 2 focal planes of the neuron with 2.5 µm separating them. The pre- and post-processed extended depth of field images acquired with just the IPP pupil mask are shown in Figs. 6.14(c) & (d) respectively. The image was deconvolved using the FISTA algorithm described in Chapter 5 with β = 3. The corresponding images for the IPP-BPM configuration are shown in Fig. 6.14. In both cases the deconvolution has significantly improved the resolution of the EDF image. The IPP-BPM image exhibits a longer DOF than the IPP mask alone, and thus several image features have been better preserved. This effect is especially noticeable in regions with high background fluorescence from features significantly beyond the depth of field. Despite
Figure 6.13: Comparison of simulated and experimental IPP-BPM PSF. The axial line profiles (a) agree well with theory, however the experimentally observed transverse profile (b) is again broader than predicted. The experimental (c) r-z PSF cross section shows additional fabrication and alignment errors when compared to simulation (d).
the decreased resolution with respect to the theoretical model, the IPP and IPP-BPM pupil masks have achieved the same DOF and resolution as a stopped down aperture without sacrificing light collection ability.

A second experiment observed bovine pulmonary artery endothelial (BPAE) cells that are labeled with Texas Red-x phalloidin emitting at 608 nm. For this experiment the BPM pupil mask etched to the central wavelength of 592 nm was used. The preparation was tilted to create an artificial DOF of approximately 8.5 \( \mu \text{m} \). This sample offers minimal background fluorescence and does not feature many specimen induced aberrations. Two clear aperture focal planes are shown in Figs. 6.15(a) and (b). Significant blurring can be seen over the FOV in both images. The pre- and post-processed IPP EDF images are shown in Figs. 6.15(c) and (d). The corresponding images for the IPP-BPM images are shown in Figs. 6.15(e) and (f) respectively. As with the Purkinje cell sample, FISTA was used to solve the deconvolution problem, however due to the lower background and noise levels the regularization parameter was set to \( \beta = 0.1 \). It is again clear that the IPP-BPM mask offers a longer DOF than the IPP mask alone, although its advantages are less drastic in this case as there is minimal to no background fluorescence.

6.4 Conclusions

The work described above demonstrates a number of firsts in the field of hybrid imaging for EDF. To the author’s knowledge it is the first experimental demonstration of an EDF imaging system that utilizes both pupil partitioning and wavefront modulation. A similar concept was theoretically proposed by Chu et al. \[87, 88\] utilizing aspheric surfaces, however it was never experimentally demonstrated. Despite the decreased performance associated with sub-optimal phase mask fabrication, important insights were gleaned. The first is that femtosecond glass micro machining techniques are still not well suited for fabricating precision optical elements, even those with relatively large feature sizes such as the IPP pupil mask. Another key insight is the concept of mask splitting that allows pupil partitioning and phase modulation to be physically realized in a single microscopy system. Finally, the good agreement of the axial intensity profiles with the theory
Figure 6.14: EDF images of a Purkinje neuron axon. The structure is too thick to be imaged simultaneously. There is a 2.5 µm separation between the first (a) and second (b) focal planes. When imaged with the IPP pupil mask (c) and deconvolving (d) the entire structure is shown in a single frame, though features near the edge of the DOF have been reconstructed poorly. Imaging with the IPP-BPM mask (e) and deconvolving (f) shows the entire structure well reconstructed.
Figure 6.15: Clear aperture images of BPAE cells with 8.5 µm separating the first (a) and second (b) focal planes. Imaging with the IPP mask (c) and deconvolving (d) the entire structure is shown a single frame, although features near the edge of the DOF have been reconstructed poorly. Imaging with the IPP-BPM mask (e) and deconvolving (f) shows the entire structure well reconstructed.
validates the approach of pupil partitioning combined with pupil modulation. With more precise fabrication techniques the performance of the system can be improved to fall more in line with the theoretical transverse resolution and previous demonstrations in the literature.
Chapter 7
Conclusions and Future Directions

This thesis presented a full systems level study of hybrid imaging to extend the depth of field of biological microscopes. A high level classification system was derived for the dozens of EDF pupil masks shown in the literature. The taxonomy identified the core properties of these pupil masks, and elucidated their effect on the PSF and MTF. The cubic, axicon, binary phase modulated (BPM), and incoherently partitioned pupil (IPP) pupil masks were selected for thorough modeling within the framework of high numerical aperture microscopy. These masks were chosen for in-depth modeling as they are representative of the literature, and have a diverse set of properties. The imaging model made minimal assumptions incorporating aspects such as polarization and polychromatic illumination. From these simulations both the detrimental and beneficial properties of each mask were investigated. The study yielded key insights that led to the design of a new type of pupil mask that incorporates the most beneficial properties from the other masks. The new pupil function was named the incoherently partitioned pupil with annular binary phase modulation (IPP-BPM).

The IPP-BPM pupil mask was simulated and shown to be well conditioned in terms of axial and radial symmetry. It was also able to produce a highly uniform intensity response through focus. These properties make the IPP-BPM particularly well suited for the task of quantitative biological imaging. It was shown to be twice as effective at extending the depth of field as the IPP mask. The IPP-BPM was quantitatively compared to the other EDF pupil functions and a clear aperture using the metrics of resolution, axial Strehl ratio, average axial intensity, and Hilbert
space angle.

The study of hybrid imaging for EDF microscopy was then extended to digital post-processing steps needed to improve the resolution, contrast, and fidelity of EDF images. For these deconvolution models it is absolutely necessary to have an accurate model for the noise present in the system. The primary noise sources relevant to microscopy systems were investigated including shot noise and the various sources of electronic detector noise. Specific noise models were formalized for the most commonly used detectors in biological imaging: the CCD, sCMOS, and EMCCD cameras.

These detailed noise models were then incorporated into a total variation regularized deconvolution model. The regularized deconvolution problems studied are state of the art processing schemes that are capable of denoising images while still maintaining high resolution features. Deconvolution models were presented for all three camera types previously mentioned. The deconvolution framework for EMCCD cameras represents the first such model specifically tailored for the unique noise properties of these cameras. Two related, but distinct, minimization algorithms were then investigated for solving the regularized inverse problem. The first method, the split-Bregman algorithm, represents one of the fastest methods for solving regularized problems of this type, but can be troublesome to utilize due to its many hyper-parameters. The fast iterative/shrinkage thresholding algorithm (FISTA) represents a slightly slower, but simpler solution method. It was found that the split-Bregman algorithm is unquestionably the superior from a speed of convergence perspective. However, less tangible factors such as ease of use are in favor of FISTA and are also an important consideration. This is especially true when the end user has minimal if any training in the mathematics of signal and image processing.

The deconvolution model was then applied to simulated EDF images of the various pupil mask previously mentioned. It was found that the IPP-BPM yielded the highest quality images after processing. This was true from a qualitative sense in terms of the artifacts present, as well as in a quantitative sense in terms of PSNR. The superior performance of the IPP-BPM mask held across a wide range of intensity levels.

Finally the designed IPP-BPM pupil mask was fabricated and tested in biological appli-
cations. For practical fabrication reasons the mask was split into two parts, an IPP mask and a BPM mask, and combined together optically. The IPP mask was fabricated using a femtosecond 3D glass etching process. While this technique is quite attractive in terms of cost and flexibility, the end product was not of sufficient optical quality. The BPM mask was fabricated using traditional lithography and glass etching techniques. The BPM mask also suffered from manufacturing errors, however these can be minimized with more time and refinement in the fabrication process. The deficiencies in the IPP mask are more fundamentally linked to the femtosecond fabrication process. The PSF of the IPP and IPP-BPM pupil masks were measured experimentally and compared with theory. After accounting for spherical aberration the axial intensity profiles of both masks were in good agreement with theory. The transverse PSF profiles were significantly broader in both cases, however this can be directly attributed to the fabrication error previously discussed. The two masks were then tested in biological imaging experiments of fluorescently labeled specimens. In both cases the masks significantly extended the DOF, and some of the lost resolution was recovered after deconvolution.

7.1 Future Directions

The field of hybrid imaging for EDF, in general, is a relatively mature technology with respect to the fundamental theory. The work presented in this thesis has helped to translate and extend this theory to the world of high resolution microscopy systems. With respect to the IPP-BPM pupil mask, there is still room for continued design and improvement through the use of stochastic optimizers. Although it is my belief that the primary areas of growth will generally be found in the fabrication and implementation of these systems. There is still a tremendous need for a reliable, efficient, and precise pupil mask manufacturing process so that the technique can be used more readily in biological imaging laboratories. A detailed tolerance analysis would aide in this goal through the identification of suitable fabrication techniques.

The work presented in this thesis has expanded hybrid imaging to applications involving quantitative experiments. There are currently plans to test the proposed system in the imaging of...
neurons in experiments attempting to address questions about the creation and storage of memories. Other planned experiments will image yeast cultures with the aim of better understanding fundamental biological process behind cell division. As the fabrication process for these masks becomes more streamlined, the ultimate goal will be to identify and solve a variety of relevant biological imaging problems that push the limits of extended depth of field microscopy.
Bibliography


