ANALYSIS OF MULTI-CHANNEL MICROSCOPY: SPECTRAL SELF-INTERFERENCE, MULTI-DETECTOR CONFOCAL AND 4PI SYSTEMS

BRYNMOR J. DAVIS

Dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

BOSTON

UNIVERSITY

BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Dissertation

ANALYSIS OF MULTI-CHANNEL MICROSCOPY: SPECTRAL

SELF-INTERFERENCE, MULTI-DETECTOR CONFOCAL AND 4PI

SYSTEMS

by

BRYNMOR J. DAVIS

B.E. (Hons I), University of Canterbury, 1999
M.S., University of Arizona, 2001

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2006

Approved by

First Reader

William C. Karl, PhD
Professor of Electrical and Computer Engineering and Professor of
Biomedical Engineering

Second Reader

Anna K. Swan, PhD
Research Associate Professor of Electrical and Computer Engineering

Third Reader

Bahaa E. A. Saleh, PhD
Professor of Electrical and Computer Engineering

Fourth Reader

Bennett B. Goldberg, PhD
Professor of Physics and Professor of Electrical and Computer Engineering

Fifth Reader

M. Selim Ünlü, PhD
Professor of Electrical and Computer Engineering

Acknowledgments

ANALYSIS OF MULTI-CHANNEL MICROSCOPY: SPECTRAL SELF-INTERFERENCE, MULTI-DETECTOR CONFOCAL AND 4PI SYSTEMS

(Order No.   )

BRYNMOR J. DAVIS

Boston University, College of Engineering, 2006

Major Professor: William C. Karl, PhD,
Professor of Electrical and Computer Engineering,
Professor of Biomedical Engineering

ABSTRACT

Fluorescence microscopy is an important and ubiquitous tool in biological imaging due to the high specificity with which fluorescent molecules can be attached to an organism and the subsequent nondestructive in-vivo imaging allowed. Focused-light microscopies allow three-dimensional fluorescence imaging but their resolution is restricted by diffraction. This effect is particularly limiting in the axial dimension as the diffraction-limited focal volume produced by a lens is more extensive along the optical axis than perpendicular to it. Approaches such as confocal microscopy and 4Pi microscopy have been developed to improve the axial resolution. Spectral Self-Interference Fluorescence Microscopy (SSFM) is another high-axial-resolution technique and is the principal subject of this dissertation. Nanometer-precision localization of a single fluorescent layer has been demonstrated using SSFM. This accuracy compares favorably with the axial resolutions given by confocal and 4Pi systems at similar operating parameters (these resolutions are approximately 350nm and 80nm respectively).

This theoretical work analyzes the expected performance of the SSFM system when imaging a general object, i.e. an arbitrary fluorophore density function rather than a single layer. An existing model of SSFM is used in simulations to characterize the system's resolution. Several statistically-based reconstruction methods are applied to show that the expected resolution for SSFM is similar to 4Pi microscopy for a general object but does give very high localization accuracy when the object is known to consist of a limited number of layers. SSFM is then analyzed in a linear systems framework and shown to have strong connections, both physically and mathematically, to a multi-channel 4Pi microscope. Fourier-domain analysis confirms that SSFM cannot be expected to outperform this multi-channel 4Pi instrument. Differences between the channels in spatial-scanning, multi-channel microscopies are then exploited to show that such instruments can operate at a sub-Nyquist scanning rate but still produce images largely free of aliasing effects. Multi-channel analysis is also used to show how light typically discarded in confocal and 4Pi systems can be collected and usefully incorporated into the measured image.

Preface

The work presented in this dissertation was carried out during my time as a graduate student at Boston University. This period was from the Fall semester of 2001 through the Spring semester of 2006. The publications directly resulting from this research and on which I am listed as an author, can be found in the bibliography and are cited below.

(Davis et al., 2005a)

(Davis et al., 2005c)

(Davis et al., 2004a)

(Davis et al., 2004b)

(Davis et al., 2004c)

(Goldberg et al., 2004)

(Swan et al., 2003)

Throughout the dissertation these papers will be referenced where the relevant research is presented. Two related provisional patents were also filed.

(Davis et al., 2005b)

(Ünlü and Davis, 2005)

Additionally, several papers on unrelated research were published during my stay at Boston University.

(Davis and Nawab, 2004)

(Davis et al., 2004d)

(Davis et al., 2002)

Contents

List of Figures

List of Abbreviations

APD	....	Avalanche Photodiode
ASF	....	Amplitude Spread Function
ATF	....	Amplitude Transfer Function
CARS	....	Coherent Anti-Stokes Raman Spectroscopy
DFT	....	Discrete Fourier Transform
DIC	....	Differential Interference Contrast
DKFZ	....	Deutsches Krebsforschungzentrum
FCS	....	Fluorescence Correlation Spectroscopy
FFT	....	Fast Fourier Transform
FIEFK	....	Fredholm Integral Equation of the First Kind
FLIM	....	Fluorescence Lifetime Imaging Microscopy
FRAP	....	Fluorescence Recovery After Photobleaching
FRET	....	Förster Resonance Energy Transfer
FTIR	....	Fourier Transform Infrared
I^5M	....	Incoherent Interference Illumination, Image Interference Microscopy
LS	....	Least Squares
ME	....	Maximum Entropy
MIAM	....	Multiple Imaging Axis Microscopy
ML	....	Maximum Likelihood
MMM	....	Multifocal Multiphoton Microscopy

Chapter 1

Introduction

1.1 Significance of this Work

The field of microscopy is broad in both the range of techniques available and the scope of applications to which it can be applied. The scale of the acquired images can vary from millimeters to nanometers depending on the application and the physical properties imaged also vary greatly. Images are formed in one, two or three spatial dimensions and when other parameters such as time or wavelength are recorded, the dimensionality is increased even further. Applications include such diverse fields as integrated-circuit inspection, material characterization and biological research.

The development of new microscopy techniques is currently an active field of research. The maturation and acceptance of any new technique is dependent on a number of milestones being reached — these include such things as a physical demonstration of the instrument, application to a relevant imaging problem(s), theoretical justification of the technique's principle of operation and a characterization of the instrument's capabilities and limitations. Spectral self-interference fluorescence microscopy is an emerging microscopy technique that shows promise in the field of biological imaging. The purpose of this work is to further the development of this technique by reaching some of the milestones mentioned above.

The resolution of a microscopy system is one of its most important parameters as it effectively defines the maximum meaningful magnification that can be achieved. This work quantifies the resolution of SSFM and compares it to other microscopy systems. A theoretical framework is also developed for SSFM that allows the effects of modifications to the

system to be understood and also gives an easy way of comparing it to other microscopy systems. This allows SSFM's capabilities and limitations to be fully understood. The theoretical framework is also used to examine schemes for improving imaging speed and reducing sensitivity to noise using SSFM. Similar analysis is also applied to novel operating methods in two well-known techniques — 4Pi microscopy and confocal microscopy. Essentially, this work develops a thorough quantitative theoretical analysis of the new spectral self-interference fluorescence microscopy technique.

1.2 General Microscopy Background

This section includes a brief and non-exhaustive introduction to microscopy (particularly fluorescence microscopy) and some of the more important techniques in the field. Fluorescence microscopy has become an indispensable tool in biological imaging due to its capacity to produce in-vivo images of a range of specimen features. It's basic principle of operation can be summarized as follows.

1. A fluorescent dye is selected from the wide range available. These fluorescent molecules are then used to stain the object to be imaged. Each particular dye attaches to a certain well-known feature in the object — this leads to highly specific imaging of that particular feature.

2. The sample is illuminated with light at the excitation wavelength of the fluorophores. This causes the fluorophores to emit light at a longer wavelength.

3. The emitted light is collected and used to form an image.

Many fluorescent dyes are compatible with living organisms which allows in-vivo imaging. Some other methods, such as electron microscopy (Egerton, 2005), X-ray microscopy (Shinohara and Yada, 1990) or Near-field Scanning Optical Microscopy (NSOM) (Pohl and Courjon, 1993) offer much higher resolution but are not compatible with live cells and are hence incapable of measuring the dynamics of a living organism. Other in-vivo methods such as absorption-, reflection- or scattering-based microscopy (Murphy, 2001), phase

contrast microscopy (Zernike, 1955), Differential Interference Contrast (DIC) microscopy (Smith, 1955; Nomarski, 1955), Second Harmonic Generation (SHG) microscopy (Gannaway and Sheppard, 1978), Coherent Anti-Stokes Raman Spectroscopy (CARS) (Begley et al., 1974; Duncan et al., 1982; Zumbusch et al., 1999) or Optical Coherence Tomography (OCT) (Huang et al., 1991; Schmitt, 1999) rely on the inherent optical properties of the sample to produce contrast in the image. These contrast mechanisms may be very weak (and hence difficult to image) or result in a low resolution image. Fluorescence microscopy allows the contrast mechanism to be introduced artificially by means of the dye. This allows great flexibility and specificity in what can be imaged. In particular, functional imaging of a biological process can be achieved by selecting a dye that attaches to a dynamic structure of interest within the sample.

The most basic form of fluorescence microscopy involves a uniform illumination of the sample so that all fluorophores are excited. The emitted light is then collected through a lens in order to form an image of the fluorophores. This is known as widefield fluorescence microscopy and is a popular method of producing a two-dimensional image of the fluorophore distribution in the focal plane of the lens. More sophisticated fluorescence microscopy methods have focused on three-dimensional imaging and improving the resolution of the measured image.

Three-dimensional imaging can be achieved using a widefield microscope and scanning the focal plane through the object. The axis of scanning, which is along the optic axis of the lens, is known as the 'axial' dimension while the plane perpendicular to this is defined by two 'lateral' axes. Axially-scanning widefield images are typically enhanced by image reconstruction algorithms to give a technique known as deconvolution microscopy (Agard and Sedat, 1983; Holmes, 1988; Agard et al., 1989). The term 'deconvolution' is used interchangeably with 'image reconstruction' in this work. The resolution limits for these widefield techniques have been known for some time (Abbe, 1873) and are due to the effects of diffraction. The resolution of the system increases with the numerical aperture (NA) of the lens and decreases with the wavelength of light used. The axial resolution is

significantly lower than the lateral resolution. As an example, a deconvolution microscopy system operating in water, at a wavelength of 500nm and with a NA of 1 has a maximum lateral resolution of about 250nm and a maximum axial resolution of about 1000nm. The method for calculating these figures is given in Chapter 5.

Confocal microscopy (Minsky, 1955; Wilson and Sheppard, 1984) has the potential to provide a significant improvement in both axial and lateral resolution when compared to deconvolution microscopy. In this method, the illumination light is focused and the emitted light is also focused on to a detector with a very limited aperture. This has the effect of producing a narrower plane of focus in the final image, which implies a higher axial resolution. For example, a confocal system operating in water, at a wavelength of 500nm and with a NA of 1 gives a maximum lateral resolution of about 125nm and a maximum axial resolution of about 500nm. However, the focused excitation does mean that only one point in the object is imaged at a time. As such, creating a three-dimensional image requires three-dimensional scanning (c.f. one-dimensional scanning in deconvolution microscopy). Due to this three-dimensional scanning, the object may be subjected to significantly more illumination light. This makes photobleaching (Tsien and Waggoner, 1990; Murray, 1998), a phenomenon in which fluorophores cease to emit after extended excitation, a significant problem.

Several techniques use varying optic-axis orientations to either reduce scanning requirements and/or provide more isotropic resolution. For example, in confocal theta microscopy (Stelzer and Lindek, 1994) illumination and detection are focused to the same point but from orthogonal directions. Multiple Imaging Axis Microscopy (MIAM) (Swoger et al., 2003) is essentially a instrument combining four widefield-deconvolution systems of differing orientation. Combining the four data sets allows the construction of an image with approximately isotropic resolution. Selective Plane Illumination Microscopy (SPIM) (Huisken et al., 2004; Huisken et al., 2005) illuminates the sample with a thin sheet of laser light and collects in widefield mode from a perpendicular orientation. This results in an improved axial resolution due to the axially confined illumination. A tomographic approach

in which the sample is rotated has also been applied in confocal microscopy (Heintzmann and Cremer, 2002).

Recent advances such as Standing Wave Fluorescence Microscopy (SWFM) (Bailey et al., 1993), Incoherent Interference Illumination, Image Interference Microscopy (I^5M) (Gustafsson et al., 1995; Gustafsson et al., 1999) and 4Pi microscopy (Hell, 1990; Hell and Stelzer, 1992a) have used interference effects to improve axial resolution. These techniques produce data that includes interference fringes of varying strength and rely on deconvolution to produce an image free of distortions. The reasons for the improved resolution are addressed in Chapter 4 and Chapter 5 but as an example, a 4Pi Type C system operating in water, at a wavelength of 500nm and with lenses of NA 1 can give resolutions of up to 125nm laterally and 100nm axially.

Other techniques such as structured illumination microscopy (Gustafsson, 2000; Frohn et al., 2001) are also used to increase lateral resolution. Non-linear fluorophore response has been used to produce significant resolution improvement in Stimulated Emission Depletion (STED) Microscopy (Hell and Wichmann, 1994; Klar and Hell, 1999; Dyba and Hell, 2002) and in several post-processing based approaches (Heintzmann and Jovin, 2002; Gustafsson, 2005). Some other techniques concentrate on particular imaging or operation scenarios. For example, two-photon microscopy (Denk et al., 1990) is well suited to imaging in turbid media; Förster Resonance Energy Transfer (FRET) imaging (Stryer, 1978; Clegg, 1996) detects when two fluorescence molecules are in very close proximity (this technique is sometimes known as Fluorescence Resonance Energy Transfer imaging); Fluorescence Lifetime Imaging Imaging Microscopy (FLIM) (Lakowicz et al., 1992; Periasamy, 2001) enables the imaging of factors that affect the longevity of the dye's fluorescent state (such as pH, oxygen and calcium ion concentration); Fluorescence Correlation Spectroscopy (FCS) (Elson and Madge, 1974; Madge et al., 1974; Thompson, 1991) and Fluorescence Recovery After Photobleaching (FRAP) (Liebman and Entine, 1974; Meyvis et al., 1999; White and Stelzer, 1999) measure fluorophore dispersion over time; and Total Internal Reflection Fluorescence (TIRF) microscopy (Axelrod et al., 1982) allows the imaging of a very thin plane

along the underside of a reflecting surface. There are many surveys, of varying extent, addressing different microscopy techniques, e.g. (Slayter and Slayter, 1992; Gustafsson, 1999; Murphy, 2001; Nagorni and Hell, 2001b; Hell, 2003; Török and Kao, 2003). These give significantly more detail than the very brief taxonomy outlined here.

1.3 SSFM Background

It has been known for several decades that the measured emissions from a fluorophore are influenced by close proximity to a reflecting surface due to interference effects (Drexhage, 1970; Drexhage, 1974). This phenomenon has been exploited to provide high-accuracy measurements of biological samples by varying the fluorophore-to-reflector distance and measuring the resulting changes in measured intensity (Lambacher and Fromherz, 1996; Braun and Fromherz, 1997; Braun and Fromherz, 1998). SSFM removes the need for physically moving the fluorophores (or mirror) by increasing the fluorophore-to-mirror spacing. This results in the optical path length changing enough across the emitted spectrum to allow varying interference effects to be observed at different emitted wavelengths (Swan et al., 2002; Swan et al., 2003; Moiseev et al., 2004a). As a result, a high-accuracy estimate (on the order of nanometers) of the fluorophore's position can be found and there is also the possibility of imaging complicated fluorophore density distributions by processing the measured spectrum. The operation of SSFM is qualitatively described in more detail in the rest of this section.

Like SWFM, I^5M and 4Pi microscopy, spectral self-interference fluorescence microscopy uses interferometric effects to improve resolution. The fluorescently labeled sample is placed a significant (in relation to the wavelength of light) distance above a reflecting surface. The fluorophores can be excited by uniform light or a specifically tailored distribution (or even a number of consecutive distributions). The fluorescent emissions are then imaged by a spectral detector. The high-performance imaging of the SSFM system depends on there being two paths to the detector for the emitted light — a direct path and one reflected from the mirror. One possible implementation of SSFM is shown in Figure 1·1.
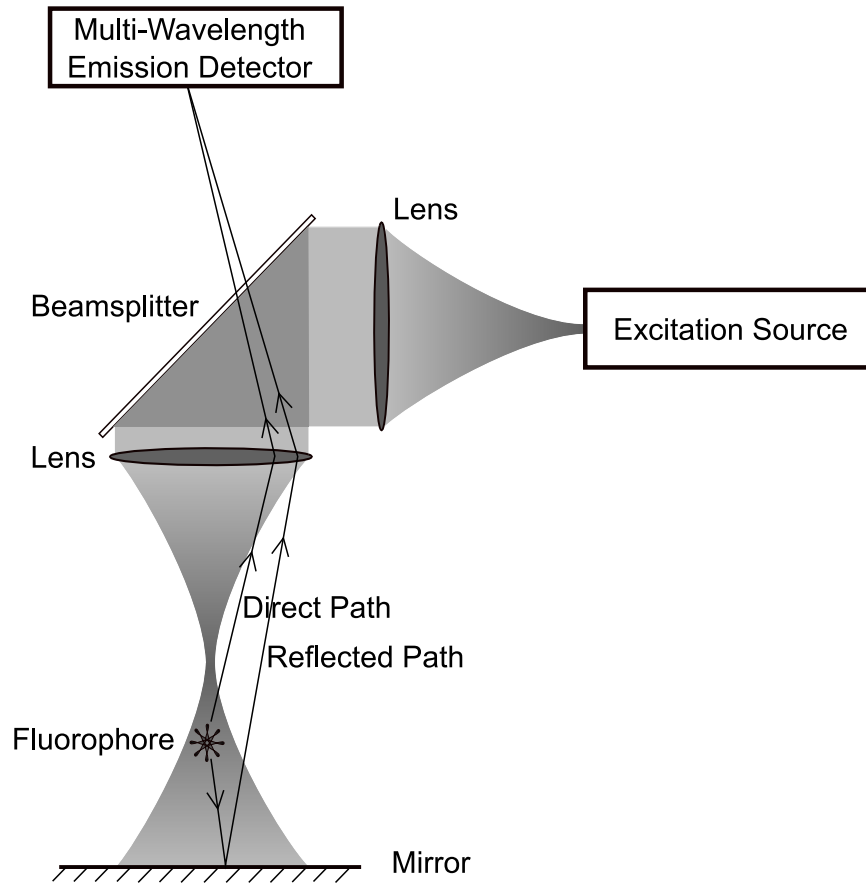
Figure 1·1: A basic illustration of one implementation of spectral self-interference fluorescence microscopy. Note that the light from the excitation source has been shown as focused while a ray illustration has been used for the emission. This is solely to provide a clear distinction between the two — both excitation and emission light will undergo focusing effects.

The light traveling along the two paths will interfere at the detector. For each wavelength detected, the effect of interference will be different, as the fixed physical distance above the mirror gives a differing optical path length. The result is that there will be oscillations in the collected spectrum as the interference cycles through constructive and destructive cases. The frequency and position of this oscillatory pattern is sensitive to the position of the fluorophore above the mirror, i.e. the distance between the mirror and the fluorophore produces a spectral signature. Analyzing the measured spectrum thus allows the fluorophore position to be estimated. An interesting comparison can be drawn between

SSFM and Fourier Transform Infrared (FTIR) Spectroscopy (Griffiths and De Haseth, 1986). The latter technique uses an interferometer to measure a spatial-domain interference pattern from an incoherent source passed through a sample. Mathematical processing is then used to reconstruct a Fourier-domain absorption spectrum for the sample. SSFM measures an Fourier-domain spectrum from an incoherent source and then uses processing to reconstruct the spatial-domain object.

It has been demonstrated (Swan et al., 2003) that if the fluorophores are confined to a single height above the mirror, that height can estimated to nanometer precision. Applications for this type of imaging have also been shown (Goldberg et al., 2004; Moiseev et al., 2004b; Moiseev et al., 2004c; Moiseev et al., 2006). The achieved localization accuracy is orders of magnitude better than the resolution of conventional techniques such as confocal microscopy and 4Pi microscopy. For these instruments the axial resolution is approximately 350nm and 80nm respectively for visible wavelengths. The sort of sensitivity demonstrated by SSFM is very promising but further analysis is required, particularly in regard to imaging general fluorophore distributions (rather than localizing a single layer) and characterizing the performance of the system. This dissertation contains such an analysis and several other related results in multi-channel microscopy.

1.4   Dissertation Outline

In Chapter 2 some general issues involving the construction of a mathematical model for the SSFM system are addressed. This includes a statistical explanation of fluorescence spectra measurements, a method of discretizing the continuous physical model and a description of the assumed noise statistics.

In Chapter 3, a test object is defined and the model used to generate simulated data. Several physically-based estimation techniques are then applied to produce a reconstruction of the test object. A specialized technique for reconstructing objects made up of a few non-zero components is also developed and tested. In all cases the expected performance of a SSFM system is shown to be similar to that of a 4Pi instrument.

This connection is made more definite in Chapter 4 where the SSFM instrument is shown to have strong connections to a multi-channel 4Pi instrument. As a result, a comprehensive system model can be constructed in terms of linear shift-invariant systems and their resulting Point Spread Functions (PSFs). This model also allows easy comparisons with other microscopy systems to be performed.

The connection between SSFM and 4Pi systems is further developed in Chapter 5 where the PSF construction is interpreted in the Fourier domain. This allows it to be shown that the generalizations that take a 4Pi instrument to a SSFM instrument do not increase the Optical Transfer Function (OTF) support. No extension in the support implies that there is no fundamental improvement in resolution. The OTF construction result also has important implications in other resolution-improvement techniques such as PSF engineering through the application of pupil filters. This application is explored in detail.

While SSFM cannot provide improved resolution over a 4Pi system, it has been shown to provide axial localization with only a single spectral measurement. By comparison, a 4Pi system must scan through the object to provide resolution. This difference is investigated in Chapter 6 where it is shown that multi-channel instruments can be used to reduce scanning requirements. The resulting lower scanning rate may provide an improved image acquisition time. To provide a comparison, the reduced-scanning, multi-channel instrument is compared to a standard-scanning, multi-channel system but with the sample-dwell time reduced to give the same image acquisition time. A noise analysis shows that the standard-scanning instrument will produce superior image quality.

In Chapter 7 a multi-channel microscope is analyzed and compared to a similar single-channel instrument. It is shown that multi-channel instruments have the potential to produce images that are less affected by noise. This relies on collecting light that would otherwise be discarded and usefully processing the additional data when forming the image.

A brief summary of the dissertation and a discussion of future work is given in Chapter 8. An appendix gives a summary of the notation before the references are listed. A curriculum vitae (as at graduation) concludes this work.

Chapter 2

SSFM Modeling Framework

This chapter describes the process involved in constructing a mathematical model for SSFM. A rigorous, continuous, deterministic model of the system has already been constructed (Swan et al., 2003). The new work presented in this chapter consists of several more subtle points in the development of a model suitable for computing solutions to the inverse problem. It is shown how the random emissions of a fluorescing molecule can statistically be expected to produce a predictable interferometric signal; the continuous, deterministic model is discretized so that it is suitable for numerical computations; a model for the noise on the observations is specified; it is shown how multiple observations can be incorporated into the model; and a short chapter summary is given to conclude. Much of the notation used in future chapters is also introduced.

2.1   Interferometric Spectral Measurements and Statistical Optics

The basic idea of spectral self-interference (as applied to fluorescence microscopy) is to estimate a fluorophore density $o(\mathbf{r})$ by interfering the emitted light with itself and examining the spectrum. At first glance, the idea that interference can be measured from an incoherent source seems somewhat counterintuitive. This section presents a rigorous mathematical explanation for this phenomenon. The analysis shows that taking spectral measurements isolates each spectral component which prevents the broad bandwidth of the fluorophore emission washing out the interference effects.

The following analysis treats light in a statistical framework as in (Goodman, 1985). In addition to accounting for the non-deterministic nature of the signals measured, the statistical approach is particularly appropriate for optics, where time averages of signals

are all that can practically be measured (due to limitations imposed by detector speeds).

Let $\dot{\xi}(t; \mathbf{r})$ be the temporal electromagnetic signal emanating from the fluorophores at position $\mathbf{r}$ (bold typeface indicates a vector). The $\dot{}$ term over $\dot{\xi}(\cdot, \cdot)$ indicates that the function is in the temporal domain. The absence of a dot indicates that the function is in the spectral domain, which is the more natural domain for this work. Note that $\dot{\xi}(t; \mathbf{r})$ is scalar and real — scalar optics will be used in this section for simplicity and complex notation will not be introduced at this point.

The emitted light is assumed to be temporally ergodic which means that the expectation (over time) shown below is equal to the equivalent time average. Time averages of functions of the field are typically what is measured by an optical detector due to detector speed limitations.

$$E_t[\dot{\xi}(t; \mathbf{r})\dot{\xi}(t - \tau; \mathbf{r} - \mathbf{r}')] \propto |h_e(\mathbf{r})|^2 o(\mathbf{r})\dot{\epsilon}(\tau)\delta(\mathbf{r}') \tag{2.1}$$

Here $h_e(\mathbf{r})$ is the excitation amplitude pattern. This term appears as a magnitude squared as the fluorophore responds to optical intensity rather than optical amplitude. The term $\dot{\epsilon}(\tau)$ is the temporal autocorrelation of the fluorophore emission or equivalently, the inverse Fourier transform of the spectral envelope $\epsilon(\nu)$.

$$\epsilon(\nu) = \mathcal{F}\{\dot{\epsilon}(\tau)\} = \int_{-\infty}^{\infty} \dot{\epsilon}(\tau)e^{-i2\pi\nu\tau}d\tau \tag{2.2}$$

From Equation 2.1 it can be seen that the autocorrelation of the emitted light is proportional to the autocorrelation of the fluorophores' emission multiplied by the excitation intensity and the density of fluorophores present. The $\delta(\mathbf{r})$ indicates that the emissions are uncorrelated across position — this corresponds to the fluorophore emissions being spatially incoherent. The temporal coherence is defined by the autocorrelation function $\dot{\epsilon}(\tau)$ or equivalently $\epsilon(\nu)$. The coherence time is typically defined as the reciprocal of the width of $\epsilon(\nu)$, which can easily be shown to result in a coherence length approximated by the equation below.

$$\text{Coherence Length} \approx \frac{\lambda_{\text{central}}^2}{\lambda_{\text{max}} - \lambda_{\text{min}}} \tag{2.3}$$

Here the width of spectrum $\epsilon(\lambda)$ (written as a function of wavelength) is given by $\lambda_{\max}-\lambda_{\min}$ and this is assumed to be relatively small in comparison to the central wavelength λ_{central}. For a fluorophore emitting at a central wavelength of 550nm with a spectral width of 100nm the resulting coherence length is approximately 3μm. This indicates an incoherent source — the field is uncorrelated with itself after propagation of about 5–6 wavelengths. As will be subsequently seen, spectral detection allows interferometric effects to be observed even when the interferometer uses path-length differences in excess of the coherence length.

It will be assumed that the instrument is linear and time-invariant. Linearity is dependent on the detector and the fluorophores not being saturated. Time-invariance simply requires that the instrument does not vary significantly over the span of a measurement. It will also be assumed that the only light collected originates from the fluorophores. This assumption requires that the instrument is well shielded from outside sources of light and/or that any spurious contributions can be subtracted from the measured signal in a preprocessing step (Shotton, 1993). These assumptions mean that the temporal signal at the detector $\dot{\chi}(t)$ is a linear function of the signals emitted from the fluorophores.

$$\dot{\chi}(t) = \int\int \dot{\xi}(t - t'; \mathbf{r})\dot{h}_d(t'; \mathbf{r})dt'd\mathbf{r} \tag{2.4}$$

Where $\dot{h}_d(t; \mathbf{r})$ is the temporal detection impulse response of the system as a function of position.

The spectral detector gives a measure of the incident intensity at each optical frequency. This will be denoted by $d(\nu)$ where ν is the temporal frequency. In statistical language,

the measured quantity is the power spectral density (Papoulis, 1984).

$$
\begin{aligned}
d(\nu) \quad &\propto \quad \mathcal{F}_\tau\{E_t[\dot{\chi}(t)\dot{\chi}(t-\tau)]\} \\
&= \quad \mathcal{F}_\tau\left\{E_t\left[\iint \dot{\xi}(t-t';\mathbf{r})\dot{h}_d(t';\mathbf{r})dt'd\mathbf{r}\iint \dot{\xi}(t-\tau-t'';\mathbf{r}')\dot{h}_d(t'';\mathbf{r}')dt''d\mathbf{r}'\right]\right\} \\
&= \quad \mathcal{F}_\tau\left\{\iiiint E_t[\dot{\xi}(t-t';\mathbf{r})\dot{\xi}(t-\tau-t'';\mathbf{r}')]\dot{h}_d(t';\mathbf{r})\dot{h}_d(t'';\mathbf{r}')d\mathbf{r}\,d\mathbf{r}'dt'dt''\right\} \\
&= \quad \mathcal{F}_\tau\left\{\iiiint |h_e(\mathbf{r})|^2 o(\mathbf{r})\dot{\epsilon}(\tau-t'+t'')\delta(\mathbf{r}-\mathbf{r}')\dot{h}_d(t';\mathbf{r})\dot{h}_d(t'';\mathbf{r}')d\mathbf{r}\,d\mathbf{r}'dt'dt''\right\} \\
&= \quad \int \mathcal{F}_\tau\left\{\iint \dot{\epsilon}(\tau-t'+t'')\dot{h}_d(t';\mathbf{r})\dot{h}_d(t'';\mathbf{r})dt'dt''\right\}|h_e(\mathbf{r})|^2 o(\mathbf{r})d\mathbf{r} \\
&= \quad \int \epsilon(\nu)|h_d(\nu;\mathbf{r})|^2|h_e(\mathbf{r})|^2 o(\mathbf{r})d\mathbf{r} \quad\quad\quad (2.5)
\end{aligned}
$$

The final step of the derivation above comes from the Fourier domain properties of convolution. Here $h_d(\nu, \mathbf{r})$ is the temporal Fourier transform of $\dot{h}_d(t;\mathbf{r})$. That is, it represents the detection system's effect on a monochromatic signal of temporal frequency ν emanating from position $\mathbf{r}$. The convention of capitalizing Fourier representations of functions will be reserved for spatial Fourier transforms which will be much more common in this work.

The final result of Equation 2.5 is an observation model for the SSFM system. This continuous model is in the form of a Fredholm Integral Equation of the First Kind (FIEFK) (Kress, 1989). This equation basically states that each position in the sample has a spectral signature given by $\epsilon(\nu)|h_d(\nu;\mathbf{r})|^2|h_e(\mathbf{r})|^2$. The data is produced by weighting each spectral signature by the corresponding fluorophore density and summing. Separating the spectral signatures in the data and recognizing their weightings allows one to determine the fluorophore density $o(\mathbf{r})$. The promise of SSFM is in that it produces spectral signatures that vary rapidly as a function of $\mathbf{r}$ due to interference effects. This means nearby points in the object produce significantly different data which allows them to be resolved.

The ideas behind the continuous model of Equation 2.5 will now be illustrated with a very simple example. Only the axial dimension z will be considered and a crude simplification of the physics of the system will be used but the gross behavior of SSFM is captured.

Assume that the sample is illuminated with a uniform field (i.e. $h_e(\mathbf{r}) = 1$) and that the received signal is related to the emitted signal as follows.

$$\dot{\chi}(t) = \int \dot{\xi}(t; z) + \rho \, \dot{\xi}(t - \tau(z); z) \, dz \tag{2.6}$$

That is, the light travels unaltered along a direct path to the receiver and experiences only a real scaling (ρ) and position dependent delay ($\tau(z)$) along the reflected path. This simple model gives the following expression for the detection system.

$$\dot{h}_d(t; z) = \delta(t) + \rho \, \delta(t - \tau(z)) \tag{2.7}$$

$$\Rightarrow |h_d(\nu; z)|^2 = 1 + \rho^2 + 2\rho \cos(\nu \tau(z)) \tag{2.8}$$

Thus each position produces a signature of $\epsilon(\nu)[1 + \rho^2 + 2\rho \cos(\nu \tau(z))]$. This says that the frequency of the data's oscillations across ν is $\tau(z)$. This means that the fluorophore density at each z value can be found by finding the strength of the data oscillation at frequency $\tau(z)$. As would be expected, the data is windowed by the spectral envelope of the fluorophore $\epsilon(\nu)$. Note that in this example increasing the delay $\tau(z)$ between the two signals only changes the frequency of the spectral oscillations. Since spectral detection is used, the interference effects do not wash out as the path difference is increased as is usually expected with an incoherent source. This is simply because each emitted wavelength is measured separately. However, the spectrometer must have a finite resolution and at some large value of path difference the interference effects will start to wash out within this resolution width.

2.2 A Discrete Model with Noise

2.2.1 Semi-Discrete Observations

The relation in Equation 2.5 takes a continuous object into a continuous data set. In reality, the data measured with the spectrometer will be a discrete set of numbers. For this reason it is desirable to have a semi-discrete model, i.e. one that relates a continuous

object to a discrete data set.

Let the $M \times 1$ vector $\mathbf{d}$ represent the discrete data collected. In addition to being discrete, this data will also be noisy. The equation below shows how the expected value of the mth term of $\mathbf{d}$ is related to the continuous data of Equation 2.5.

$$
\begin{aligned}
E[(\mathbf{d})_m] &= \int_{\Delta_\nu^{(m)}} d(\nu) \, d\nu \\
&\approx d((\boldsymbol{\nu})_m) \int_{\Delta_\nu^{(m)}} d\nu \\
&\propto d((\boldsymbol{\nu})_m)
\end{aligned}
\tag{2.9}
$$

In this work a subscript outside a bracketed vector indicates a specific component of the vector, e.g. $(\mathbf{d})_m$ is the mth component of $\mathbf{d}$. The term $\Delta_\nu^{(m)}$ in Equation 2.9 is the integration window over which $d(\nu)$ contributes to $(\mathbf{d})_m$. Physically, this corresponds to the spectral resolution of the spectrometer. For the approximation in Equation 2.9 to hold, this resolution must be finer than the scale at which the continuous data varies — i.e. if the spectral resolution is too low, the interference fringes will wash out and the instrument model will not be valid. The bracketed superscript indicates that $\Delta_\nu^{(m)}$ is the mth in a series of Δ_ν terms and in this case Δ_ν partitions the ν axis. The vector $\boldsymbol{\nu}$ contains the midpoints of the windows i.e. sampling points. The midpoint rule (Delves and Mohamed, 1985) is used to discretize the observations in Equation 2.9 and a constant of proportionality is discarded. These methods simply mean that the discrete data will be modeled by sampling the continuous data of Equation 2.5.

$$
E[(\mathbf{d})_m] = \int \epsilon((\boldsymbol{\nu})_m) |h_d((\boldsymbol{\nu})_m; \mathbf{r})|^2 |h_e(\mathbf{r})|^2 o(\mathbf{r}) \, d\mathbf{r}
\tag{2.10}
$$

2.2.2 Modeling Noise

The expected value of the data is given in Equation 2.10 but for the purposes of numerical simulation and probabilistic inversion methods, it is necessary to fully define the data's statistical properties. While uncertainty in microscopical data comes from several sources

(Sheppard et al., 1992; Pawley, 1994), it will be assumed that the noise from sources such as detector dark current, illumination fluctuations and thermal electronic noise is small in comparison to the photon counting noise. This assumption is helpful in that it gives an instrument-independent noise model, which is Poisson as it is based on photon-counting statistics. Mathematically, this means the probability of observing a given count (say b) at a given spectral data point can be expressed as shown below.

$$
\begin{aligned}
P\left\{(\mathbf{d})_m = b\right\} &= \mathcal{P}\{b; E[(\mathbf{d})_m]\} \\
&= \frac{E[(\mathbf{d})_m]^b}{b\,!} e^{-E[(\mathbf{d})_m]}
\end{aligned}
\tag{2.11}
$$

Here $\mathcal{P}\{\cdot\,;\cdot\}$ is the Poisson distribution — the first argument is the outcome and the second argument is the Poisson distribution's mean. It is also assumed that any two data points are statistically independent.

2.2.3   Fully Discretizing the Model

While Equation 2.10 represents the reality of taking a continuous object to a discrete data set, it is often desirable to have a fully discrete model, i.e. one that takes a discrete approximation of the object to a discrete data set. A fully discrete model is particularly useful for performing computational simulations. A fully discrete model can also be used to generate numerical inversion methods for application to the collected data. As long as the object is discretized on a sufficiently fine scale, the fully discrete model should be a good approximation to the semi-discrete case.

The model can be fully discretized using a method similar to that employed in Equa-

tion 2.9.

$$
\begin{aligned}
E[(\mathbf{d})_m] &= \int \epsilon((\boldsymbol{\nu})_m)|h_d((\boldsymbol{\nu})_m;\mathbf{r})|^2|h_e(\mathbf{r})|^2 o(\mathbf{r})\,d\mathbf{r} \\
&= \sum_{q=1}^{Q} \int_{\Delta_{\mathbf{r}}^{(q)}} \epsilon((\boldsymbol{\nu})_m)|h_d((\boldsymbol{\nu})_m;\mathbf{r})|^2|h_e(\mathbf{r})|^2 o(\mathbf{r})\,d\mathbf{r} \\
&\approx \sum_{q=1}^{Q} \epsilon((\boldsymbol{\nu})_m)|h_d((\boldsymbol{\nu})_m;(\boldsymbol{\zeta})_q)|^2|h_e((\boldsymbol{\zeta})_q)|^2 o((\boldsymbol{\zeta})_q) \int_{\Delta_{\mathbf{r}}^{(q)}} d\mathbf{r} \\
&\propto \sum_{q=1}^{Q} \epsilon((\boldsymbol{\nu})_m)|h_d((\boldsymbol{\nu})_m;(\boldsymbol{\zeta})_q)|^2|h_e((\boldsymbol{\zeta})_q)|^2 o((\boldsymbol{\zeta})_q) \quad (2.12)
\end{aligned}
$$

The spatial volume occupied by the object was partitioned into a set of voxels denoted by $\Delta_{\mathbf{r}}$ and $\boldsymbol{\zeta}$ is a vector of the voxel centers. The voxels are ordered so that they are indexable by a single number q, as is standard in discrete image reconstruction (Andrews and Hunt, 1977). The method of this ordering does not effect the underlying mathematics.

Since Equation 2.12 is now fully discrete, it can be written as a matrix equation.

$$
E[(\mathbf{d})] = \mathrm{diag}(\boldsymbol{\epsilon})\,C\,\mathbf{o} \quad (2.13)
$$

The notation $\mathrm{diag}(\cdot)$ indicates that the elements of the vector argument are placed down the main diagonal of a square matrix that is otherwise zeros. The data vector $\mathbf{d}$ was defined earlier (Equation 2.9) and the remaining terms above are defined as follows.

$$
(\mathbf{o})_q = o((\boldsymbol{\zeta})_q) \qquad\qquad (Q \times 1) \qquad\qquad (2.14)
$$

$$
(\boldsymbol{\epsilon})_m = \epsilon((\boldsymbol{\nu})_m) \qquad\qquad (M \times 1) \qquad\qquad (2.15)
$$

$$
(C)_{mq} = |h_d((\boldsymbol{\nu})_m;(\boldsymbol{\zeta})_q)|^2|h_e((\boldsymbol{\zeta})_q)|^2 \qquad\qquad (M \times Q) \qquad\qquad (2.16)
$$

The equations above show that this method involves simply sampling the functions of interest. The matrix C and the envelope $\boldsymbol{\epsilon}$ are kept separate in Equation 2.13 because the spectral envelope $\boldsymbol{\epsilon}$ may not be precisely known. By contrast, C can be calculated. It may be necessary to estimate $\boldsymbol{\epsilon}$ in the process of constructing an estimate of the object.

2.2.4   Including Multiple Excitation and Detection Patterns

Up to this point, the models described have related a single collected spectrum to the object. However, multiple sets of data may be collected by using different excitation and/or detection patterns. This allows additional information about the object to be measured.

Say J different excitation/detection patterns are used to give J sets of data. That is $\mathbf{d}^{(1)}, \ldots, \mathbf{d}^{(J)}$ are collected with corresponding models $C^{(1)}, \ldots, C^{(J)}$. The object vector ($\mathbf{o}$) remains the same for each measurement and the fluorophore-emission spectral envelope is also constant. These multiple observation models can be written in a single augmented form.

$$E[(\mathbf{d})] = \mathcal{E}\, C\, \mathbf{o} \tag{2.17}$$

The variables $\mathbf{d}$ and C from the previous section are augmented (as shown below) when multiple observations are taken.

$$\mathbf{d} = \begin{bmatrix} \mathbf{d}^{(1)} \\ \vdots \\ \mathbf{d}^{(J)} \end{bmatrix} \qquad (JM \times 1) \tag{2.18}$$

$$C = \begin{bmatrix} C^{(1)} \\ \vdots \\ C^{(J)} \end{bmatrix} \qquad (JM \times Q) \tag{2.19}$$

The term $\mathcal{E}$ is a diagonal matrix formed by repeating the envelope vector J times and represents the effect of the fluorophore-emission spectral envelope on multiple observations.

$$\mathcal{E} = \mathrm{diag}\left(\begin{bmatrix} \boldsymbol{\epsilon} \\ \vdots \\ \boldsymbol{\epsilon} \end{bmatrix} \right) \qquad (JM \times JM) \tag{2.20}$$

The analysis above shows how multiple observations can also be combined into a single matrix model. The single- and multiple-observation expressions in Equation 2.13 and

Equation 2.17 are in the same form and are both suitable for the use in the application of numerical inversion methods.

2.3 Summary

This section has shown how a deterministic, continuous, physical model can be used to create a discrete (or semi-discrete) stochastic model. The physical model provided in (Swan et al., 2003) was created by analyzing the instrument at each frequency ν emitted by the fluorophore. Section 2.1 shows how this is valid even though the fluorophore emits a random, broadband signal. In Section 2.2 the discrete nature of the observations are accounted for, noise is modeled and a fully discrete version of the model is given. In short, this chapter has laid the groundwork for creating a model that can be used to generate simulated data (both noisy and noise-free), to evaluate system performance through simulation and to develop inversion methods for collected data.

Chapter 3

SSFM Image Reconstruction and Performance Evaluation

This chapter contains the results of applying several different object estimation approaches to the SSFM inverse problem. These results are found through computer simulation — that is, the fully discrete model is used to generate synthetic data from a given object, noise is added and an inversion method (commonly called deconvolution in the microscopy community) is applied to get an estimate of the object. Comparing the original object to the estimated object gives an idea of the performance that can be expected from the actual instrument, provided that the model is well-matched to physical reality. Since SSFM is primarily a technique for enhancing axial resolution, the examples in this chapter will be one-dimensional — i.e. the object is assumed to vary slowly laterally which means it is sufficient to display only axial profiles. In this chapter and throughout this dissertation the term 'resolution' refers to the minimum feature size visible in an arbitrary object while 'localization' refers to the accuracy achievable when estimating the position of a feature known to be present in the object. As will be seen, this is an important distinction.

In addition to explaining and applying several inversion methods to SSFM, this chapter contains analysis of several issues more specifically related to SSFM — an upper limit to the number of useful excitation patterns is derived and a method of estimating the fluorophores' spectral emission envelope is given. The inversion methods are also applied to several other light microscopy systems in order to get a performance comparison. Preliminary results in evanescently-excited SSFM are also presented.

3.1 Determining the Number of Useful Excitation Patterns

Due to the deliberately induced reflections in the SSFM system, the excitation light will exhibit interferometric structure through the object. In the next chapter a comprehensive model will be developed that includes factors such as focusing effects when modeling the excitation field. However at this point it will be assumed that the depth of focus of the excitation light is long enough to allow the following standing-wave approximation for the excitation intensity.

$$|h_e(z)|^2 = 1 + |\rho_e|^2 + 2\text{Re}\{\rho_e\} \cos\left(\frac{4\pi n}{\lambda_e} z + \phi_e\right) \tag{3.1}$$

Here ρ_e is a complex reflection coefficient, n is the index of refraction, λ_e is the excitation wavelength and ϕ_e captures any phase shift present (e.g. at reflection and/or deliberately introduced). Assuming that the user has some mechanism for varying ϕ_e, the object can be excited by patterns $|h_e^{(1)}(z)|^2, |h_e^{(2)}(z)|^2, \ldots, |h_e^{(M)}(z)|^2$ corresponding to phases $\phi_e^{(1)}, \phi_e^{(2)}, \ldots, \phi_e^{(M)}$. At this point it is important to ask how much additional information is provided with each new excitation pattern.

The expression in Equation 3.1 is just a constant plus a sinusoidal term and is thus determined (to within a scaling factor) by two parameters (e.g. the phase of the sinusoidal component and the amplitude of the constant component). It is a simple exercise in algebra to show that any three $|h_e(z)|^2$ terms with different phases can be linearly combined to produce any given profile in the form of Equation 3.1. This means that anything more than the first three excitation patterns are redundant in the sense that the data measured could also be constructed from a simple operation on the previous three data sets. This result is entirely analogous to one presented in standing wave microscopy (Krishnamurthi et al., 1996).

Throughout this chapter three equally spaced excitation phases will be used in the simulations.

$$\phi_e^{(1)} = 0, \qquad \phi_e^{(2)} = \frac{2\pi}{3}, \qquad \phi_e^{(3)} = \frac{-2\pi}{3} \tag{3.2}$$

Spacing the phases equally over the full 2π range ensures that there is not a preferential phase direction.

It could be argued that while more than three excitation patterns doesn't give any additional data in the noise-free case, it would allow improvements in noisy data quality by using noise averaging techniques. This idea will not be pursued as the additional time required to take extra measurements could instead be used to collect for a longer time (and hence increase the number of measured photons) in each of the first three data sets. This would improve the noisy data quality also.

3.2   General Inversion Approaches

Image reconstruction and inverse problems are a well-studied field in general (Bertero, 1989; Hansen, 1998; Karl, 2000) and many approaches have been proposed specifically for microscopy applications (Holmes, 1988; Holmes, 1989; Bertero et al., 1990; Carrington et al., 1995; Conchello and McNally, 1996; van Kempen et al., 1997; Verveer et al., 1999). In this section the fundamental issues in image reconstruction will be identified and explained, and several appropriate techniques applied. This is by no means a representation of all inversion approaches that could be applied to the SSFM problem.

The object used in these simulations is shown in Figure 3·1. The spacing between the data points is 2nm which, as will be seen later, is very fine compared to the resolution of the microscopy systems of interest. This means the simulated data generated from the object should be very close to what would be expected from a continuous-space (i.e. semi-discrete) measurement. The object was chosen to have a variety of spacings between the non-zero fluorophore densities and also have an extended region with non-zero density. The goal is to use an object which will test the resolving capabilities of the systems examined.

The simulated data resulting from this object is shown in Figure 3·2. The SSFM instrument is that described in (Swan et al., 2003) and the more accurate physical model from that publication is used both to generate the data and in the subsequent inversion algorithms. The three data sets corresponding to the three excitation phases are shown,
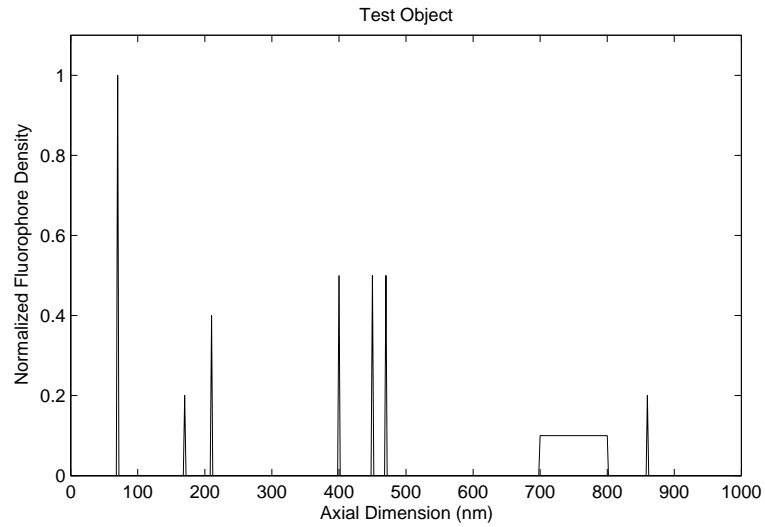
Figure 3·1: The test object used for the numerical simulations in this chapter.
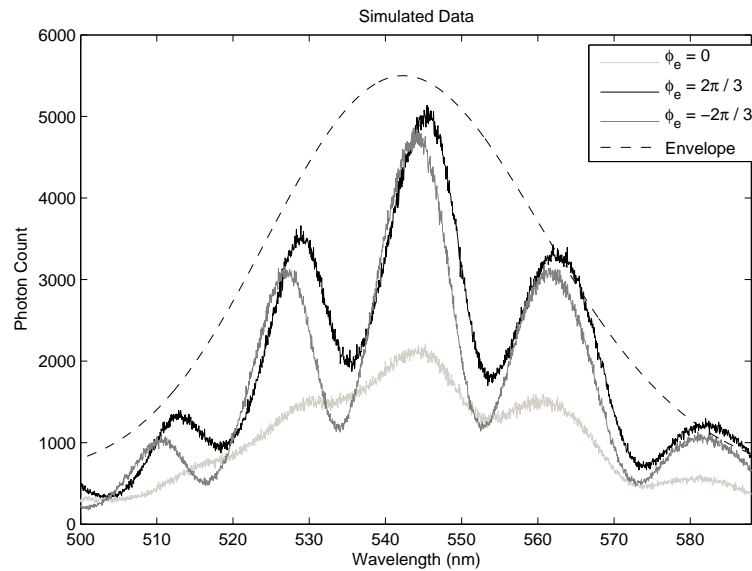
Figure 3·2: Simulated data resulting from the test object. The measurement includes three phases of the standing-wave excitation and the peak expected count is 5000. The spectral emission envelope used in the model is also plotted.

along with the assumed fluorophore-emission spectral envelope. The reflectance used in the standing-wave excitation model was 0.5 and the object was placed 5μm above the reflecting surface. The signal level is consistent with measurements taken in the Optical Characterization and Nanophotonics Laboratory at Boston University.

For the first part of this chapter it will be assumed that the spectral envelope ϵ is known exactly. This means that the model of the SSFM system is also known and can be used in the deconvolution process.

$$E[\mathbf{d}] = \mathcal{E}C\mathbf{o} = \mathcal{C}\mathbf{o} \tag{3.3}$$

The matrix $\mathcal{C}$ can therefore be calculated. The inversion procedure can be modified to deal with an unknown envelope, as described later in the chapter. The collected data $\mathbf{d}$ is a random vector modeled by Equation 3.3 and a Poisson noise model. These facts will be used to define statistically-based inversion procedures.

3.2.1 Maximum Likelihood Estimation

Maximum Likelihood (ML) estimation consists of choosing the object estimate according to the rule below.

$$\hat{\mathbf{o}} = \underset{\mathbf{o}}{\mathrm{argmax}} \ \log\left[P(\mathbf{d}|\mathbf{o})\right] \tag{3.4}$$

That is, the object estimate is chosen such that it maximizes the probability of the observed data conditioned on the object estimate. ML estimation is a very popular technique due to its optimality with respect to the Cramer-Rao bound (Kay, 1993) for unbiased estimators. Maximum likelihood estimation is also commonly applied in optics and microscopy (Frieden, 1972; Holmes, 1988; Holmes, 1989; Markham and Conchello, 2001).

Using the Poisson distribution from Equation 2.11 gives the following expression for

the ML estimator.

$$
\begin{aligned}
\hat{\mathbf{o}} &= \underset{\mathbf{o}}{\operatorname{argmax}} \log \left[\prod_{u=1}^{JM} \frac{(\mathcal{C}\mathbf{o})_u^{(\mathbf{d})_u}}{(\mathbf{d})_u!} e^{-(\mathcal{C}\mathbf{o})_u} \right] \\
&= \underset{\mathbf{o}}{\operatorname{argmax}} \sum_{u=1}^{JM} (\mathbf{d})_u \log[(\mathcal{C}\mathbf{o})_u] - (\mathcal{C}\mathbf{o})_u
\end{aligned}
\tag{3.5}
$$

The maximum above can be found by taking the derivative with respect to $\mathbf{o}_v$ and setting to 0.

$$
\begin{aligned}
0 &= \sum_{u=1}^{JM} (\mathbf{d})_u \frac{(\mathcal{C})_{uv}}{(\mathcal{C}\hat{\mathbf{o}})_u} - (\mathcal{C})_{uv} \\
&= \sum_{u=1}^{JM} \frac{(\mathcal{C})_{uv}}{(\mathcal{C}\hat{\mathbf{o}})_u} [(\mathbf{d})_u - (\mathcal{C}\hat{\mathbf{o}})_u]
\end{aligned}
\tag{3.6}
$$

A matrix equation can be used to represent this expression for all v.

$$
\mathcal{C}^T \Lambda^{-1}(\hat{\mathbf{o}}) \, \mathcal{C}\hat{\mathbf{o}} = \mathcal{C}^T \Lambda^{-1}(\hat{\mathbf{o}}) \, \mathbf{d}
\tag{3.7}
$$

Where the term $\Lambda(\hat{\mathbf{o}})$ is as defined below.

$$
\Lambda(\hat{\mathbf{o}}) = \operatorname{diag}(\mathcal{C}\hat{\mathbf{o}})
\tag{3.8}
$$

Here $\mathcal{C}\hat{\mathbf{o}}$ is the data expected from $\hat{\mathbf{o}}$. This determines $\Lambda(\hat{\mathbf{o}})$ which is the covariance matrix of the data expected from $\hat{\mathbf{o}}$.

The non-linear problem defined by Equation 3.7 and Equation 3.8 can be solved using a range of techniques including the Richardson-Lucy algorithm which is applied later in this chapter. However, to gain insight into the problem, an approximation will first be described. The matrix $\Lambda(\hat{\mathbf{o}})$ has the expected value of the data $\mathbf{d}$ along its main diagonal. While this vector is not known (as the object is unknown) it can be approximated by low-pass filtering the data. The filtering is applied to remove the frequency components of the noise that are higher than the expected-data frequencies.

$$
\Lambda(\hat{\mathbf{o}}) = \operatorname{diag}\{E[\mathbf{d}]\} \approx \operatorname{diag}\{\mathrm{LPF}[\mathbf{d}]\} \equiv \Lambda(\mathbf{d})
\tag{3.9}
$$

This approximation can be expected to produce acceptable results as long as the expected data is higher than a few counts and the spectrometer samples the spectrum at a sampling rate that is high in comparison to the structure of the data. Low-pass filtering to remove noise is a standard digital signal processing operation and the issues involved are well understood (Oppenheim and Schafer, 1989). For this reason the low-pass filtering operation will not be addressed further, except to note that when this estimation is performed for the data of Figure 3·2 the maximum difference between $\Lambda(\hat{\mathbf{o}})$ and $\Lambda(\mathbf{d})$ is less than 4% of $\Lambda(\mathbf{o})$.

Once the approximation of Equation 3.9 is applied, $\Lambda(\hat{\mathbf{o}})$ becomes $\Lambda(\mathbf{d})$ which is simply a constant matrix. This means that Equation 3.7 is now approximated as follows.

$$\mathcal{C}^T\Lambda^{-1}(\mathbf{d})\ \mathcal{C}\hat{\mathbf{o}} = \mathcal{C}^T\Lambda^{-1}(\mathbf{d})\ \mathbf{d} \tag{3.10}$$

This equation is equivalent to a standard weighted Least Squares (LS) error problem where $\Lambda^{-0.5}(\mathbf{d})$ is the weighting.

$$\hat{\mathbf{o}} = \underset{\mathbf{o}}{\operatorname{argmin}} \left\| [\Lambda(\mathbf{d})]^{-0.5}(\mathbf{d} - \mathcal{C}\mathbf{o}) \right\|^2 \tag{3.11}$$

Where $\|\cdot\|$ is the 2-norm. This is also equivalent to ML estimation when a Gaussian noise model, with a covariance matrix $\Lambda(\mathbf{d})$, is used. This is reasonable as a Poisson distribution tends toward a Gaussian distribution (with the same mean and variance) as the Poisson mean increases (Papoulis, 1984).

At this point, a reconstruction is not feasible as Equation 3.11 is ill-posed. Generally, the operator $\mathcal{C}$ will have a non-zero nullspace and as the reconstruction space is increased (e.g. by finer discretization in the reconstruction) a nullspace will certainly be present. This means that some object components do not affect the data. Thus Equation 3.11 will not have a unique solution and the Cramer-Rao bound will not exist. ML methods are optimal amongst the set of unbiased estimators when the Cramer-Rao bound exists but when it does not there are no performance guarantees.
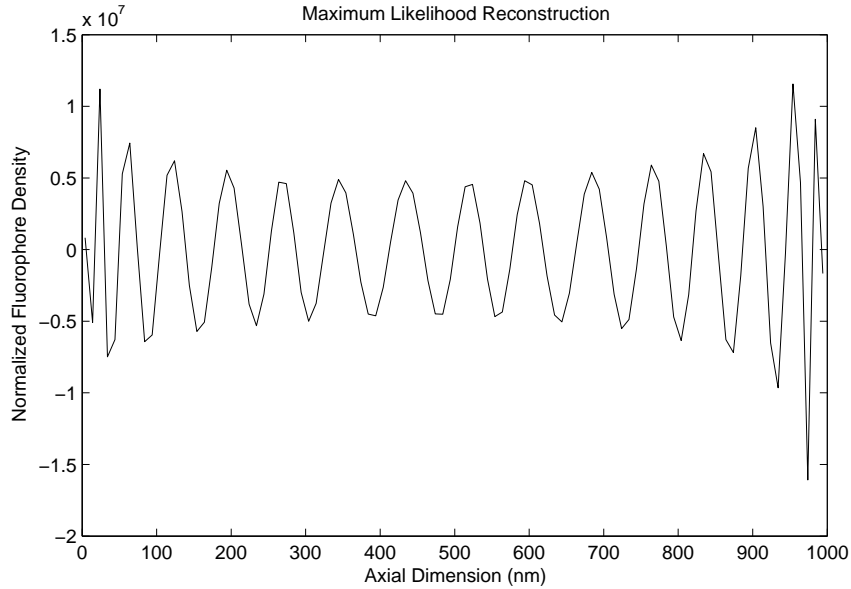
Figure 3·3: A pseudo-inverse reconstruction of the object from the simulated data. The reconstruction is highly oscillatory and has an erroneously high amplitude (the peak value should be 1) due to noise amplification.

Even if $\mathcal{C}$ has an empty nullspace, it is likely to be unstable in the presence of noise (Hansen, 1998; Karl, 2000). This is because the ML estimator will apply high gains to noisy components of the data and the result will be a reconstruction that is swamped by amplified noise. Such a reconstruction is shown in Figure 3·3, where the object was estimated using a numerical implementation of the psuedo-inverse (Bertero, 1989; Hansen, 1998) and the data from Figure 3·2. The pseudo-inverse finds the minimum-norm solution to Equation 3.11. This inversion method produces a unique solution but the result is still not usable due to noise amplification. To overcome these problems, some kind of regularization must be employed. Regularization is used to ensure a unique solution that is stable in the presence of noise.

3.2.2   Tikhonov Regularization

One popular regularization method is Tikhonov regularization (Tikhonov, 1963a; Tikhonov, 1963b; Bertero, 1989) and variations of it have been applied in microscopy, e.g. (Carrington

et al., 1995). This method is cast as a minimization problem. The cost function of Equation 3.11 is used with an additional term that penalizes the energy in the reconstruction.

$$\hat{\mathbf{o}} = \underset{\mathbf{o}}{\operatorname{argmin}} \left\| [\Lambda(\mathbf{d})]^{-0.5}(\mathbf{d} - \mathcal{C}\mathbf{o}) \right\|^2 + \alpha^2 \left\| \mathbf{o} \right\|^2 \qquad (3.12)$$

The extra term ensures that the minimization produces a unique estimate of the object. The regularization parameter α is used to trade off the two components of the cost function and controls the amount of regularization. If α is large, only strongly passed components of the object are reconstructed which results in a potentially lower-resolution image but good robustness to noise. As α is decreased a higher resolution image may be produced however noise instability will be more noticeable. In the limit, as α goes to zero, Tikhonov regularization becomes the pseudo-inverse. Tikhonov regularization introduces bias into the object estimator (and thus the Cramer-Rao bound does not apply) in order to improve stability. The estimate that satisfies Equation 3.12 can be written as shown below.

$$\hat{\mathbf{o}} = [\mathcal{C}^T \Lambda^{-1}(\mathbf{d}) \, \mathcal{C} + \alpha^2 I]^{-1} \mathcal{C}^T \Lambda^{-1}(\mathbf{d}) \, \mathbf{d} \qquad (3.13)$$

The regularization term in Equation 3.12 ensures that the inverse matrix in the equation above exists.

The Tikhonov reconstruction was calculated for several values of the regularization parameter as shown in Figure 3·4. The data points in the reconstruction are 10nm apart (this reconstruction-grid spacing will be used for all reconstructions in this section) while the test object was defined on a 2nm grid. An 'ideal reconstruction' is shown in each plot — this is formed by representing the test object on a 10nm grid. Each point on the 10nm grid is calculated as the sum of the five corresponding 2nm grid points. This is consistent with the fully-discrete observation model of Equation 2.12.

A recurring theme can be seen in the reconstructions of Figure 3·4. For low values of the reconstruction parameter, the energy of the estimate is only weakly penalized which means that highly oscillatory behavior may occur in an effort to account for the noise on
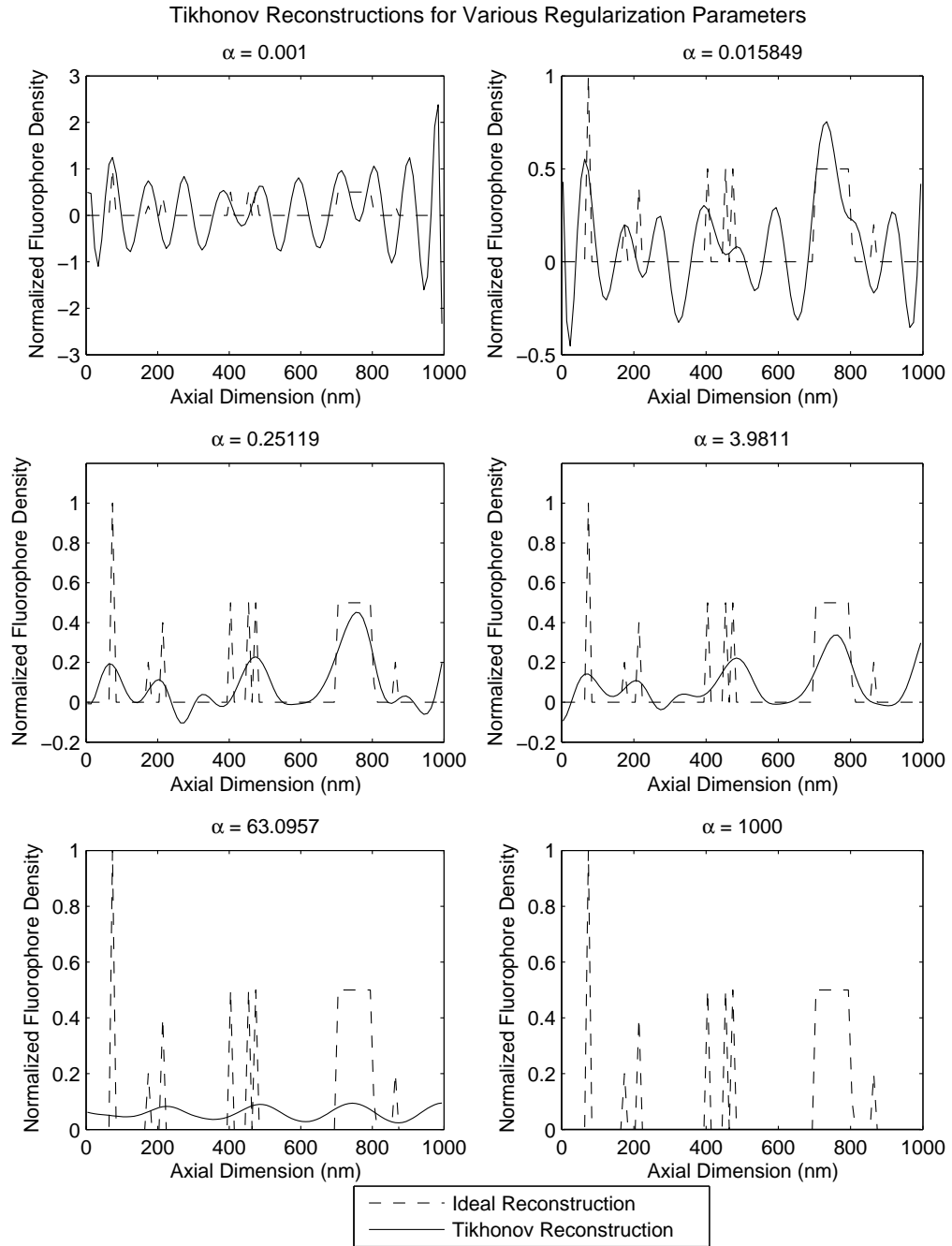
Figure 3·4: Tikhonov estimates of the test object, from the simulated data, for a variety of regularization parameters. An ideal reconstruction is also shown in each plot.
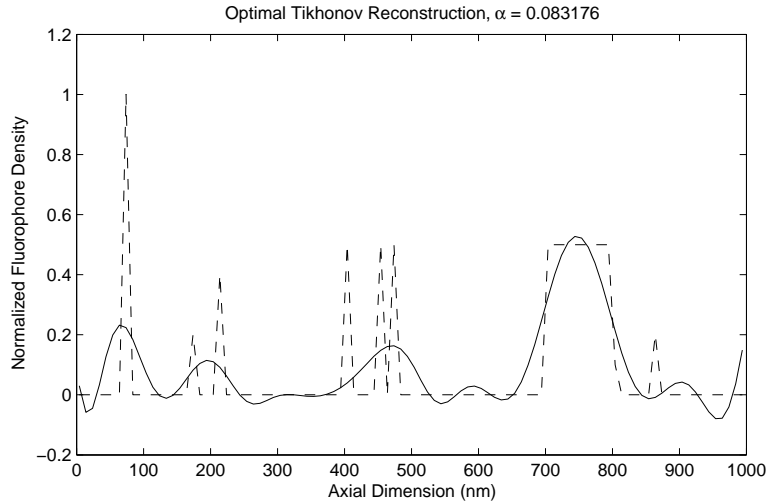
Figure 3·5: The Tikhonov reconstruction that has the least square difference from the ideal reconstruction. The ideal reconstruction is also plotted (dashed line).

the data — i.e. noise amplification occurs as in Figure 3·3. As the regularization parameter increases, the quality of the reconstruction improves up to a point. As α becomes very large the energy is penalized to such a high level that the estimate approaches zero. Finding the optimal value of α is a well known issue in regularized inverse problems.

The square error of the reconstruction ($\|\mathbf{o} - \hat{\mathbf{o}}\|$) was calculated for each regularization parameter used. The optimal reconstruction parameter, in terms of the error power, was found by finely stepping through a range of α values and evaluating the error. The best reconstruction is shown in Figure 3·5. Note that in practice the regularization parameter can not be chosen in this manner as $\mathbf{o}$ is unknown. A variety of methods exist for selecting α (Golub et al., 1979; Hansen, 1992) but will not be addressed here. Note that none of the Tikhonov reconstructions resolve the closely spaced points in the object.

3.2.3   Non-Negative Tikhonov Regularization

The reconstructions of Figure 3·4 can be seen to have negative values. This is certainly not desirable for this application as a fluorophore density function is being imaged and this must always be non-negative. Non-negative prior knowledge is enforced in a number

of applications (Snyder et al., 1992) including microscopy, e.g. in (Homem et al., 2002) non-negativity is enforced using the Projection Onto Convex Sets (POCS) algorithm — an iterative procedure applied to imaging by Gerchberg (Gerchberg, 1974) and Papoulis (Papoulis, 1975). The Tikhonov procedure can also be modified so that negative values are not produced in the reconstruction.

To achieve non-negative reconstructions, the Tikhonov definition of Equation 3.12 can be modified so that more than simply reconstruction energy is penalized by the regularizing term.

$$\hat{\mathbf{o}} = \operatorname*{argmin}_{\mathbf{o}} \left\| [\Lambda(\mathbf{d})]^{-0.5}(\mathbf{d} - \mathcal{C}\mathbf{o}) \right\|^2 + \alpha^2 \Psi_{\mathrm{pos}}(\mathbf{o}) \tag{3.14}$$

The cost function $\Psi_{\mathrm{pos}}(\mathbf{o})$ will be defined so that any negative reconstruction values will be heavily penalized.

$$\Psi_{\mathrm{pos}}(\mathbf{o}) = \sum_{q=1}^{Q} (\mathbf{o})_q^2 [1 + \alpha_{\mathrm{pos}}^2 \mathcal{H}(-(\mathbf{o})_q)] \tag{3.15}$$

The term α_{pos} should be large and $\mathcal{H}(\cdot)$ is the Heaviside step function.

This minimization problem can be solved by taking the derivative and setting it to zero. This results in the following equation which is analogous to the ML equation of Equation 3.10 and the Tikhonov solution of Equation 3.13.

$$[\mathcal{C}^T \Lambda^{-1}(\mathbf{d}) \, \mathcal{C} + \alpha^2 \Lambda_{\mathrm{pos}}^{-1}(\hat{\mathbf{o}})]\hat{\mathbf{o}} = \mathcal{C}^T \Lambda^{-1}(\mathbf{d}) \, \mathbf{d} \tag{3.16}$$

Here $\Lambda_{\mathrm{pos}}(\hat{\mathbf{o}})$ is a matrix that is zero off the main diagonal and has the following definition on the main diagonal.

$$[\Lambda_{\mathrm{pos}}(\hat{\mathbf{o}})]_{qq}^{-1} = \frac{\frac{d}{d(\hat{\mathbf{o}})_q}\Psi_{\mathrm{pos}}(\hat{\mathbf{o}})}{2(\hat{\mathbf{o}})_q} = \begin{cases} 1 & (\hat{\mathbf{o}})_q \geq 0 \\ 1 + \alpha_{\mathrm{pos}}^2 & (\hat{\mathbf{o}})_q < 0 \end{cases} \tag{3.17}$$

This result can be seen to be a weighting that heavily penalizes any negative terms. While there is an issue of existence of the derivative $\frac{d}{d(\hat{\mathbf{o}})_q}$ at $\hat{\mathbf{o}}_q = 0$, this problem has been ignored as weighting has no effect when $\hat{\mathbf{o}}_q = 0$.

The estimate defined by Equation 3.16 and Equation 3.17 is no longer easily calculated

in closed form due to the dependence of the weighting on the object estimate. However, the cost function can easily be shown to be convex which means it has a unique global minimum with no local minima (Strang, 1986). This is very helpful as it means that if a solution is found to Equation 3.16 then it is the global minimum required.

Finding this solution will be done using an iterative method. Starting from an initial guess at $\hat{\mathbf{o}}$ the weighting matrix can be calculated according to Equation 3.17. This weighting can then be used in Equation 3.16 to get a new estimate of the object. The algorithm will iterate between finding the object estimate and the weighting matrix until convergence is reached, at which point the solution to Equation 3.16 and Equation 3.17 has been found. This is a standard fixed point iteration method as used in (Charbonnier et al., 1997).

Several non-negative Tikhonov reconstructions are shown in Figure 3·6. The regularization parameter is varied over a similar range as the Tikhonov simulations shown earlier and the positivity term is weighted by $\alpha_{\mathrm{pos}} = 1 \times 10^6$. The iteration was started from the Tikhonov reconstruction (negative values allowed) with the same α value. The iterative algorithm doesn't converge for α values much smaller than those shown in Figure 3·6 (the convergence criterium was that two successive estimates don't differ by an energy more than 0.1% of the energy of the earlier estimate). As expected, negative reconstruction values are negligibly small but in general the resolution is not greatly improved.

The optimal reconstruction (in terms of square error) is shown in Figure 3·7. Again, the more closely spaced features in the object are not resolved. It is interesting to note that some of the reconstructions in Figure 3·6 may appear better than the 'optimal' one shown in Figure 3·7 as some of the closer features appear to be resolved. However, these reconstructions also tend to contain erroneous features which make them worse in terms of the chosen error metric.

3.2.4   Richardson-Lucy Reconstruction

Up to this point the reconstructions presented have used the approximation of Equation 3.9 and therefore are not strictly based on the Poissonian data statistics. The Richardson-Lucy
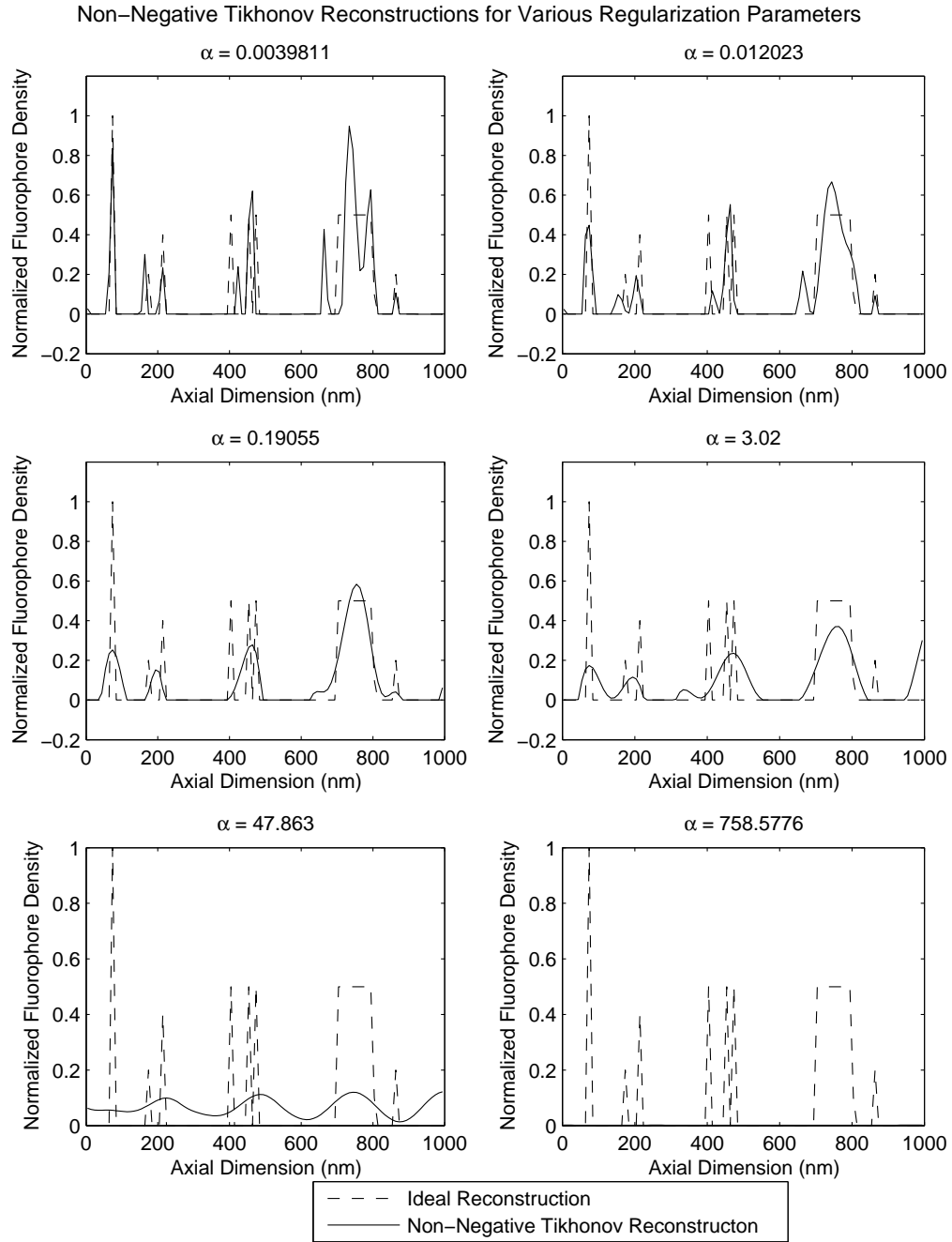
Figure 3·6: Non-negative Tikhonov estimates of the test object, from the simulated data, for a variety of regularization parameters. An ideal reconstruction is also shown in each plot.
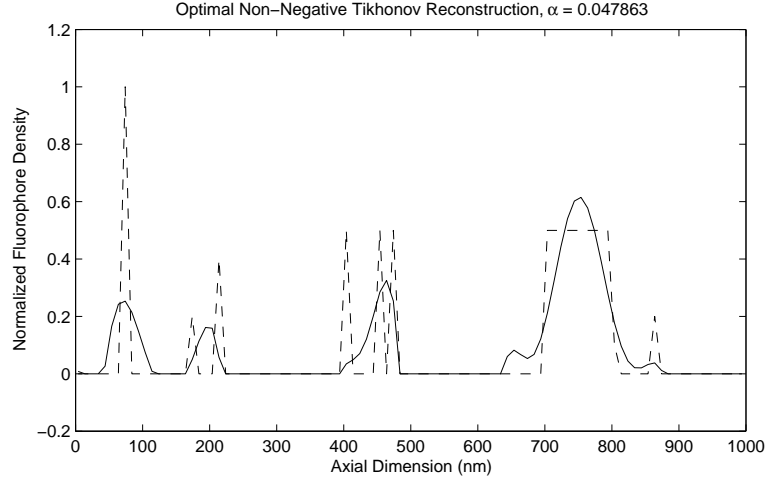
Figure 3·7: The Non-negative Tikhonov reconstruction that has the least square difference from the ideal reconstruction. The ideal reconstruction is also plotted (dashed line).

algorithm (Richardson, 1972; Lucy, 1974) is an iterative reconstruction algorithm that is matched to Poisson statistics. While it was originally developed for the reconstruction of astronomical data, it has become popular in microscopy (van Kempen et al., 1997; Nagorni and Hell, 1998; Nagorni and Hell, 2001a) where the detection model also involves a Poisson noise process.

The equation below shows one iteration of the Richardson-Lucy algorithm. This step takes the vth estimate $\hat{\mathbf{o}}^{(v)}$ and takes it to $\hat{\mathbf{o}}^{(v+1)}$.

$$\hat{\mathbf{o}}^{(v+1)} = \text{diag}(\hat{\mathbf{o}}^{(v)}) \; \mathcal{C}^T \left[\mathbf{d} \; \Lambda^{-1}(\hat{\mathbf{o}}^{(v)})\right] \left[\text{diag}(\mathcal{C}^T \mathbf{1})\right]^{-1} \tag{3.18}$$

$\Lambda(\hat{\mathbf{o}}^{(v)})$ is as in Equation 3.8 but now it is defined by the data expected from the last iteration's object estimate. Thus $[\mathbf{d} \; \Lambda^{-1}(\hat{\mathbf{o}}^{(v)})]$ is the actual data divided by the expected data. The adjoint operator $\mathcal{C}^T$ is applied to that term and the result point-by-point multiplied with the previous estimate. The term $\mathbf{1}$ is a vector of ones and $[\text{diag}(\mathcal{C}^T \mathbf{1})]^{-1}$ is used to normalize the estimate at each iteration.

The Richardson-Lucy algorithm converges to the ML estimate (for Poisson statistics) but this will have the same problems with uniqueness and stability mentioned earlier. These
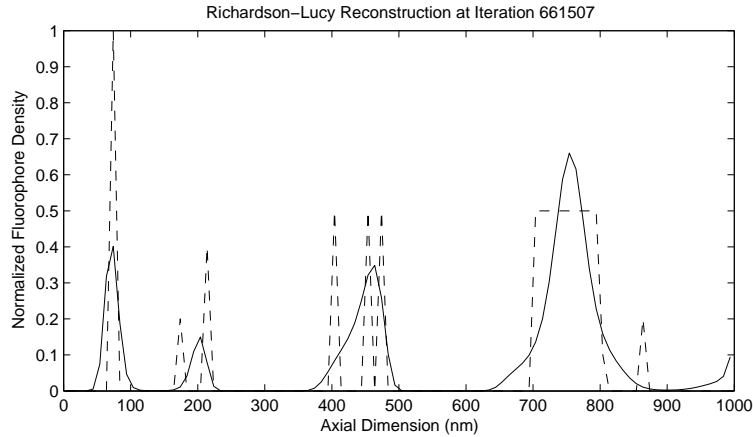
Figure 3·8: The iteration of the Richardson-Lucy algorithm that has the least square difference from the ideal reconstruction. The ideal reconstruction is also plotted (dashed line).

problems are addressed with regularization which in this case takes the form of terminating the algorithm before convergence. The earlier the algorithm is stopped, the more severe the effects of the regularization. Regularization again means that the estimate will be biased but stable. It should also be noted that since the Richardson-Lucy algorithm is matched to Poisson statistics, it will never produce negative values in the reconstruction, provided that the initial object ($\hat{\mathbf{o}}^{(0)}$) is non-negative.

The Richardson-Lucy algorithm was started from a uniform reconstruction ($\hat{\mathbf{o}}^{(0)} = \mathbf{1}$) and terminated at the iteration that gave the least square difference from the ideal reconstruction. This reconstruction is shown in Figure 3·8. The best reconstruction was reached at iteration 661507 which illustrates that the Richardson-Lucy algorithm is slow to converge. There exist algorithms to accelerate the convergence (Biggs and Andrews, 1997) but they have not been explored here. Also note that the stopping criterium used here is not applicable in real data where **o** is unknown. More practical stopping criteria for iterative algorithms have been proposed (Coakley, 1991). Qualitatively, the RL reconstruction performs similarly to the non-negative Tikhonov method and the closely spaced features are not resolved.

3.2.5 Regularization to Encourage Sparsity

In Section 3.2.3, the cost function associated with regularization was modified in order to produce a non-negative solution. This is an example of using regularization to enforce prior knowledge of the object. This idea can be used to enforce (or encourage) properties other than non-negativity. In this section, a method is used that encourages 'sparseness' of the solution. The idea is that if the object is known to consist of only a limited number of non-zero regions, this knowledge can be used to improve the reconstruction algorithm by designing it such that it preferentially produces these 'sparse' solutions (Donoho et al., 1992). In this example, a sparse reconstruction algorithm would hopefully make the blurred response around the isolated fluorophores sharper and perhaps even resolve closely spaced fluorophores. In the test object used here only 18% of its length has non-zero fluorophore density so it is arguably sparse.

In Equation 3.12 the 2-norm was used for the regularization cost. An alternative to this is to use the p-norm (Geman and Reynolds, 1992), which is defined below.

$$\|\mathbf{o}\|_p = \left(\sum_{q=1}^{Q} |(\mathbf{o})_q|^p\right)^{\frac{1}{p}} \approx \left(\sum_{q=1}^{Q} ((\mathbf{o})_q^2 + \beta)^{\frac{p}{2}}\right)^{\frac{1}{p}} \tag{3.19}$$

As p decreases from 2, the p-norm lessens for sparse vectors. The β term is added to make the derivative of $\|\cdot\|_p$ continuous and should be very small. Continuity of the derivative is desired so that calculus-based results can be used to help find the minimum of the cost function.

In addition to encouraging sparseness, it is desirable to keep the non-negativity constraint. This means Equation 3.15 and Equation 3.17 become the following when $\|\mathbf{o}\|_p^p$ is used in the regularization cost.

$$\Psi_{\text{pos}}(\mathbf{o}) = \sum_{q=1}^{Q} \left\{ [(\mathbf{o})_q^2 + \beta]^{\frac{p}{2}} + \alpha_{\text{pos}}^2 \mathcal{H}(-(\mathbf{o})_q)(\mathbf{o})_q^2 \right\} \tag{3.20}$$

$$[\Lambda_{\text{pos}}(\hat{\mathbf{o}})]_{qq}^{-1} = \begin{cases} \frac{p}{2}[(\hat{\mathbf{o}})_q^2 + \beta]^{\frac{p}{2}-1} & (\hat{\mathbf{o}})_q \geq 0 \\ \frac{p}{2}[(\hat{\mathbf{o}})_q^2 + \beta]^{\frac{p}{2}-1} + \alpha_{\text{pos}}^2 & (\hat{\mathbf{o}})_q < 0 \end{cases} \qquad (3.21)$$

The half-quadratic iterative algorithm applied in the last section can now be applied to solve this problem.

Some illustrative p-norm reconstructions are shown in Figure 3·9. The α and p parameters are varied while $\alpha_{\text{pos}} = 1 \times 10^6$ and $\beta = 1 \times 10^{-4}$ for all the reconstructions. Again, convergence was assumed when the relative energy change between iterations was less that 0.1%. In general, as the p value is decreased, sparsity is more strongly encouraged. This behavior is observed as the $\alpha = 0.8$ reconstructions tend to have fewer non-zero points than the $\alpha = 1.2$ reconstructions. Arguably this results in better reconstructions for the point-like features but the reconstructions are clearly poorer over the extended feature between z=700nm and z=800nm. Again, it can be seen that the closely spaced features are not reliably resolved without sacrificing the quality of reconstruction for extended features.

For p-norm reconstructions, local-minima issues may occur. For $p \geq 1$ it can be easily shown that the cost function is convex which implies a single minima exists and that the iterative algorithm will converge to it. For $p < 1$ convexity cannot be assured and the algorithm may converge to a local minimum. For this reason the cost function for the p=0.8 reconstructions of Figure 3·9 will be examined. The cost for the $\alpha = 0.1, 1, 10$ reconstructions are 2.1459, 2.1463 and 2.1678 respectively. If the object estimate was the ideal reconstruction shown in Figure 3·9 then the costs would be 2.9729, 2.9732 and 2.9978. These are all higher costs than for the reconstructions found. While this doesn't guarantee that the iteration isn't stuck in a local minima, it does mean that the minimization formulation is set up so that the ideal reconstruction is not optimal in terms of cost. So while local minima may be an issue, a more pressing problem is that the cost function does not force the reconstruction to the true object. It was hoped that prior knowledge about the sparsity of the object could be used to overcome the nullspace problems of the observation but this estimation formulation has not achieved that goal.

Since local minima may exist in the $p < 1$ case, the initialization point of the algorithm
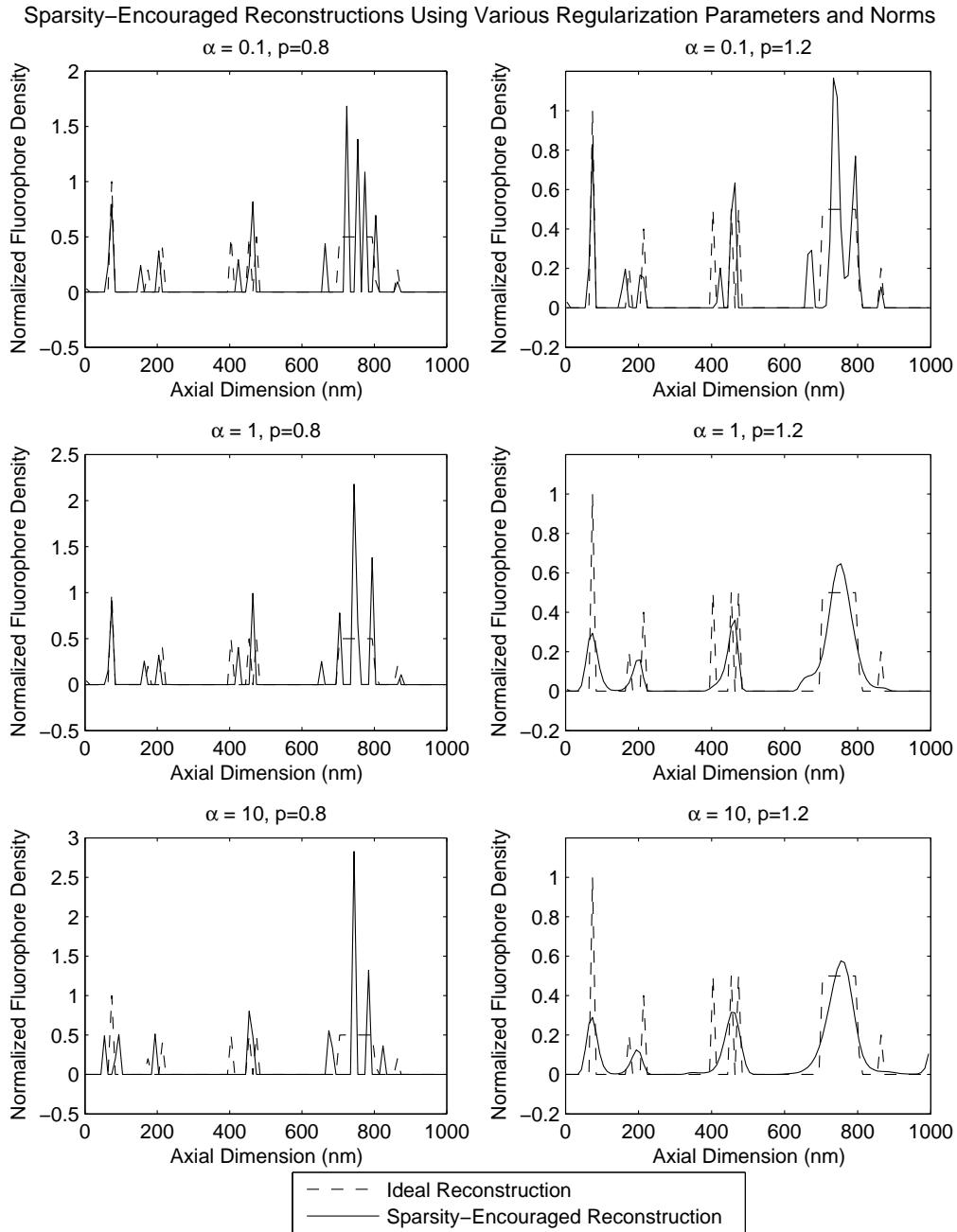
Figure 3·9: Reconstructions of the test object, from the simulated data, using a p-norm regularization term. A variety of p values and regularization parameters are used. An ideal reconstruction is also shown in each plot.
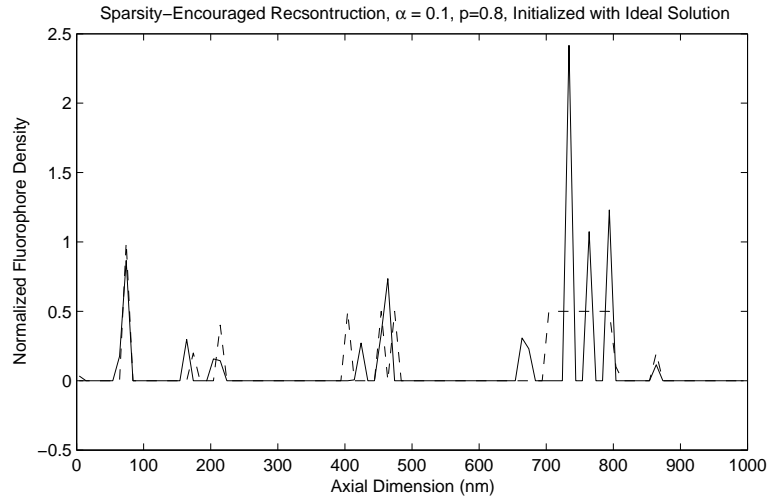
Figure 3·10: Reconstruction of the test object, from the simulated data, for a 0.8-norm regularization function and an ideal initialization point. The ideal reconstruction is also plotted (dashed line).

will be important since it will influence which minimum is found. For the reconstructions of Figure 3·9 the algorithm was started at the Tikhonov reconstruction for the same α value. The initialization point for the reconstruction in Figure 3·10 was the ideal reconstruction. The resulting reconstruction emphasizes the fact that the cost function results in the algorithm moving away from the ideal solution. It can also be seen that while the solution looks similar to its counterpart in Figure 3·9, it does in fact correspond to a different minimum — the extended object is modeled by three peaks in Figure 3·10 and four peaks in Figure 3·9. The quality of reconstruction is similar in either case.

3.3 Strong Prior Knowledge — Limiting the Number of Object Components

The simulations shown so far in this chapter have shown a resolution far from the nanometer precision that SSFM seemed to promise. This section explains the early high-precision results and how they relate to the simulation results shown above. It is shown how locating a single fluorophore layer is a very different task from trying to estimate a general distribution. The fact that an object is a single point or even a small number of known

'components' is very strong prior knowledge and can be used to give a high precision estimate of the object.

3.3.1   Posing the Limited-Component Problem

It will be assumed that the object is made up of a limited number of components. These components need not be limited to point objects but could be lines, sheets or any other known shape expected to be present in the object. An estimate of the object will be made up from a specified number of these components. This means the inversion procedure will essentially consist of picking the best model for the object out of a limited set of possibilities.

Let $\varphi^{(q)}(\mathbf{r})$ be one of Q possible unit-energy components of the object $\mathbf{o}(\mathbf{r})$. It will be assumed that the object is made up of not more than L of these components and let B be the indices of these L components. Mathematically, B is a subset of the numbers 1 through Q ($B \subseteq \{1, 2, \ldots, Q\}$) and the number of elements in B is L ($|B| = L$). Let $\mathbf{o}$ be the $L \times 1$ vector of coefficients of the basis functions in the object. This leads to the following representation of the object.

$$o(\mathbf{r}) = \sum_{q \in B} (\mathbf{o})_q \varphi^{(q)}(\mathbf{r}) \tag{3.22}$$

This representation limits the object to not more than L components (if the number of components actually present in the object is less than L some of the coefficients will be zero). Up until now the reconstructions have been a special case where $L = Q$ (i.e. any number of the components could be present) and the basis functions $\varphi^{(q)}(\mathbf{r})$ were indicator functions that were non-zero on the pixel volume $\Delta_{\mathbf{r}}^{(q)}$. For the limited-component model the observation equation becomes the following.

$$E[\mathbf{d}] = \mathcal{C}(B)\mathbf{o} = \mathcal{E}C(B)\mathbf{o} \tag{3.23}$$

The qth column of the operator $\mathcal{C}$ was the data produced from pixel q. Similarly, the limited-component operator $\mathcal{C}(B)$ has columns that are the data produced from each basis

function $\varphi^{(q)}(\mathbf{r})$. However, this operator will now vary with the L basis functions selected and will be denoted by $\mathcal{C}(B)$. Previously all the basis functions could contribute so that there was only one possible B. The forward model of Equation 3.23 is analogous to the observation model of Equation 2.16 where C was defined by Equation 2.12. In this limited component formalism the following defines $C(B)$.

$$[C(B)]_{ml} = \epsilon((\boldsymbol{\nu})_m) \int |h_d((\boldsymbol{\nu})_m; \mathbf{r})|^2 |h_e(\mathbf{r})|^2 \varphi^{(l)}(\mathbf{r}) \, d\mathbf{r}, \quad l \in B \qquad (M \times L) \quad (3.24)$$

Note that the earlier basis allowed the approximation of just sampling the excitation-detection pattern product. The more general basis requires a more complicated term — a weighted integral of the excitation-detection pattern product.

3.3.2 Solving the Limited-Component Problem

The task is now to get an estimate of B and $\mathbf{o}$ which will determine the object. Prior knowledge about the object is built in to the choice of basis functions. For example, the single layer experiments performed early in the SSFM project (Swan et al., 2003) implicitly assumed a basis of axially-offset lateral sheets. The inverse problem can again be solved by posing it as a minimization problem.

$$\hat{\mathbf{o}}, \hat{B} = \underset{\mathbf{o}, B}{\operatorname{argmin}} \left\| [\Lambda(\mathbf{d})]^{-0.5} [\mathbf{d} - \mathcal{C}(B)\mathbf{o}] \right\|^2 + \alpha^2 \Psi_{\mathrm{pos}}(\mathbf{o}) \qquad (3.25)$$

The regularization term is now just the non-negative prior from Equation 3.15 but over the set of L coefficients.

$$\Psi_{\mathrm{pos}}(\mathbf{o}) = \sum_{l=1}^{L} (\mathbf{o})_l^2 [1 + \alpha_{\mathrm{pos}}^2 \mathcal{H}(-(\mathbf{o})_l)] \qquad (3.26)$$

Restricting the object to a set of limited components will typically be a strong enough regularization procedure so that issues of noise amplification and solution-uniqueness will not arise.

With B fixed, estimating $\mathbf{o}$ is the same procedure as before. The complexity arises in

that B must also be estimated. The approach taken to this problem is somewhat inelegant and computationally expensive but guarantees an optimal solution. The space of possible B will be searched exhaustively. That is, for each possible B the optimal $\mathbf{o}$ will be found and a cost term calculated. Picking the lowest overall cost term defines both $\hat{B}$ and $\hat{\mathbf{o}}$ and thus solves the inverse problem. Of obvious concern is the computational complexity of this method which rises quickly. The computational load goes as $\frac{Q!}{L!(Q-L)!}$ (Q choose L). However, the computation is quite tractable for a few object components (small L) and can be used to give an idea of SSFM's ability to locate a single object or separate two closely spaced objects.

The method presented here is closely related to the p-norm reconstruction when $p = 0$ and this problem has received some interest recently (Donoho and Huo, 2001; Wohlberg, 2003; Malioutov, 2003). In this approach the operator $\mathcal{C}$ has a column for each of the Q basis functions (rather than using an L-column operator that varies with the components assumed present) and only a limited number of non-zero components are allowed in the object vector $\mathbf{o}$. The 0-norm simply counts the number of non-zero points so this inversion process is essentially the same as that described in Section 3.2.5 but using the 0-norm.

3.3.3 Simulation Results

Some numerical simulations will be shown to illustrate the performance of the limited-component approach. Two very simple cases will be considered — a single lateral layer and two closely spaced lateral layers. In both cases the signal level will be varied to show how the reconstruction accuracy changes with noise level.

The single layer was placed at z=100nm and multiple simulations were run, each having a different noise realization. For each simulation the layer position was estimated to the nearest 0.1nm. The results of these simulations, for three signal levels, are shown as histograms in Figure 3·11. It can be seen that with a peak expected count of 5000 photons, the estimate accuracy is within 0.1nm. Even with a peak expected count of only 50 photons the estimate accuracy is approximately 1nm. The density of the layers is also estimated
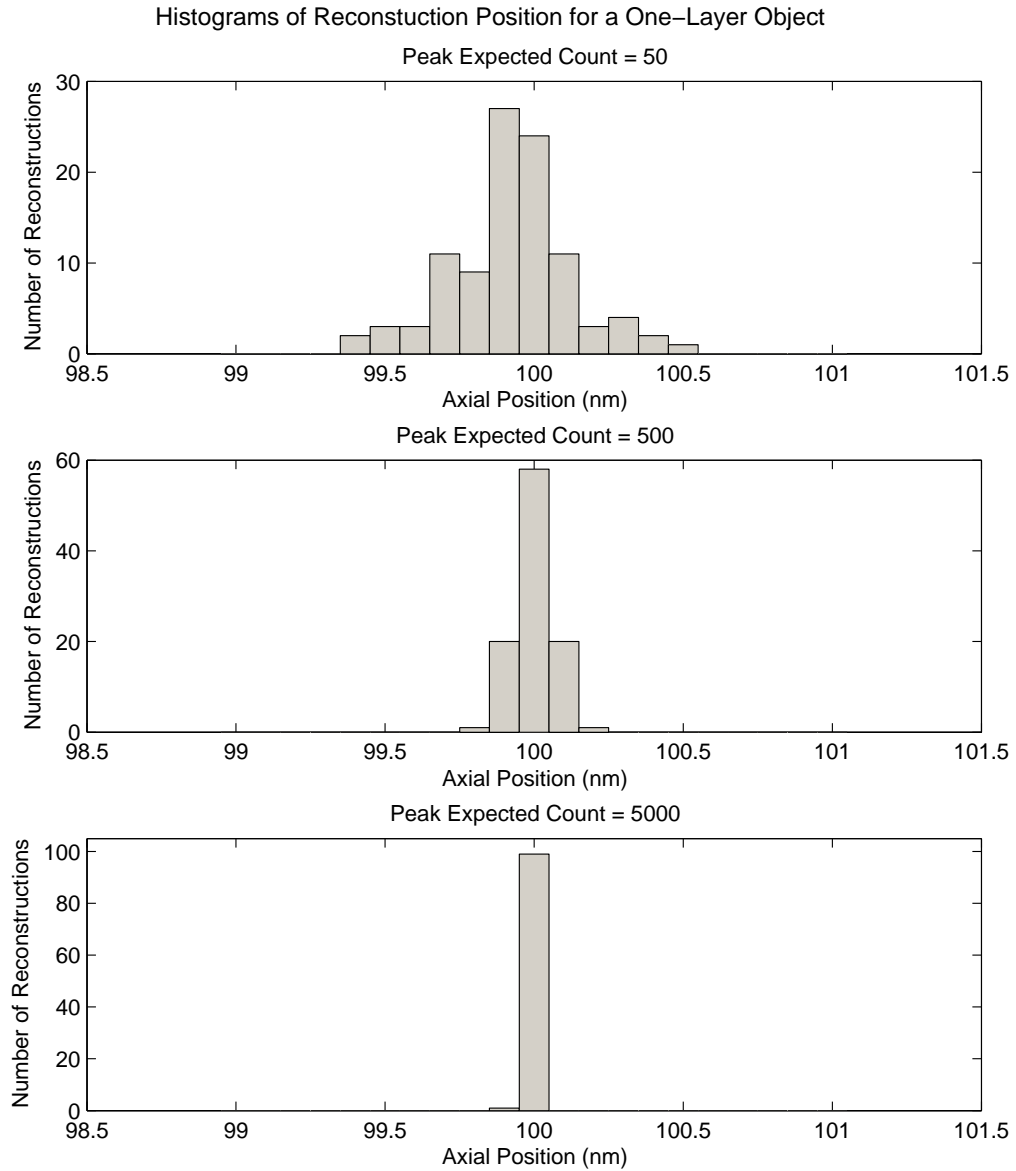
Figure 3·11: Histograms of the estimated positions of a single layer located at $z = 100$nm. The simulations are for peak expected signal levels of 50, 500 and 5000 photons. In each case 100 simulations were run, each with a different realization of the Poisson noise.

in this procedure and the standard deviations for the 50, 500 and 5000 photon cases were 0.40, 0.13 and 0.034 percent of true value in each case. These results show how a single layer can be localized much more accurately than an arbitrary object can be resolved. The high-signal simulations presented here represent a system similar to that demonstrated in (Swan et al., 2003) where nanometer-precision localization was shown experimentally.

A two-layer test object was defined with equal-density layers at z=100nm and z=120nm. A 20nm spacing is well below the achievable resolution for arbitrary objects in SSFM, 4Pi microscopy or confocal microscopy. Limited-component simulations were again run for multiple realizations of the noise random variable and for three different signal levels. The layer positions were estimated to the nearest nanometer and the results are shown as two-dimensional histograms in Figure 3·12. Each point in the plane of these images represents a pair of layer positions and the intensity represents the number of simulations that return those position estimates.

It can be seen that the localization accuracy with two layers is significantly less than with a single layer. For the 5000-count data each position estimate can vary by a few nanometers, the 500-count data shows up to 10nm variation and the 50-count data produces estimates that go right out to the edge of the positions tested (from 80nm to 140nm). The shape of the histograms indicates that the two position estimates never fall on the same side of the midpoint of the layers (in this case 110nm). This can be seen by noting the lower layer estimate does not go above 110nm and the upper does not fall below 110nm. It is also interesting to note that the error is correlated between the two layer-position estimates. This is shown by the fact that the histogram is elongated diagonally. This means that if one layer position is estimated too high, the other will most likely also be estimated as too high. Similarly both may be estimated as too low but it seldom occurs that one layer position is estimated high and the other low. Interestingly, around the true $(120, 100)$ point, the slope of the curves is approximately 1. The implication is that a significant number of reconstructions return position values that are the true values but offset by a small amount — e.g. $(119, 99)$ and $(121, 101)$ are common position estimates

Two–Dimensional Histograms of Reconstruction Positions for a Two–Layer Object
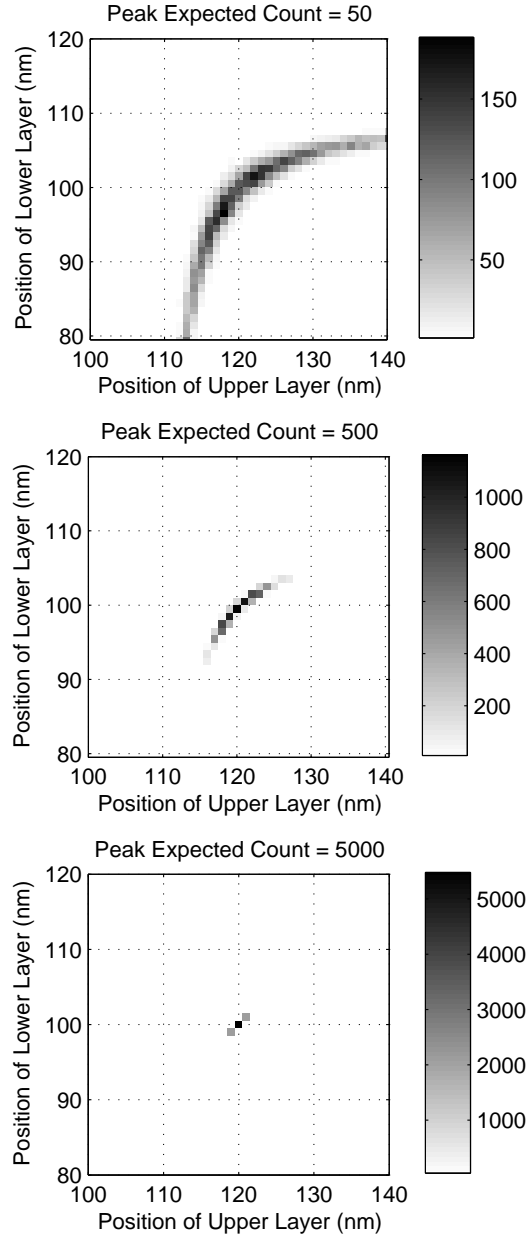
Figure 3·12: Two-dimensional histograms of the estimated positions of the layers in a two-layer object. The layers are located at $z = 100$nm and $z = 120$nm. Simulations were run for peak expected counts of 50, 500 and 5000 photons. In each case 10000 simulations were run, each with a different realization of the Poisson noise.

but $(121, 99)$, $(120, 99)$, $(121, 100)$, $(119, 100)$, $(120, 101)$ and $(119, 101)$ do not occur as frequently. This shows the spacing between the layers is estimated more robustly than the absolute positions. The estimated layer spacings for these simulations are plotted in Figure 3·13. This data was found by simply taking the difference of the two position estimates in each simulation run. The histograms show that the layer spacing is estimated to nanometer precision in the 5000-count data and to within a few nanometers in the 500-count data. While the layer spacing is estimated well, it is important to note that the relative densities of the estimated layers have not yet been explored. For this reason care should be taken when making any claims about the ability of SSFM to estimate layer spacing.

The two-layer reconstructions return four parameters to define the object. The layer positions are determined by the two-element set B, the basis functions used to describe the object and the layer fluorophore densities are determined by the 2×1 vector $\mathbf{o}$, the elements of which determine the coefficients for the basis functions according to Equation 3.22. While the statistics of this four-dimensional reconstruction are not easily visualized, the density estimates contained in $\mathbf{o}$ do exhibit some interesting behavior that is worth investigating.

In Figure 3·14 the two position estimates are plotted on the same x axis and the two density estimates are plotted on the same y axis in a two-dimensional histogram. Each simulation run produces two position estimates and two corresponding density estimates so the histograms represent 20000 layer estimates (10000 reconstructions with 2 layers in each) and 20000 density estimates. This representation of the data does not show explicitly show the correlation between the two position estimates in a given simulation as is seen in Figure 3·12, nor is the correlation between the two density estimates shown. However, as discussed earlier, each simulation has one position below 110nm and one above — this allows the two layers to be distinguished in the plots. Note that the apparent discontinuities in the contours (especially visible in the 5000-count results) are a result of pixelization in the position estimate — the position was estimated to the nearest nanometer while the density was quantized only by computational accuracy limits.
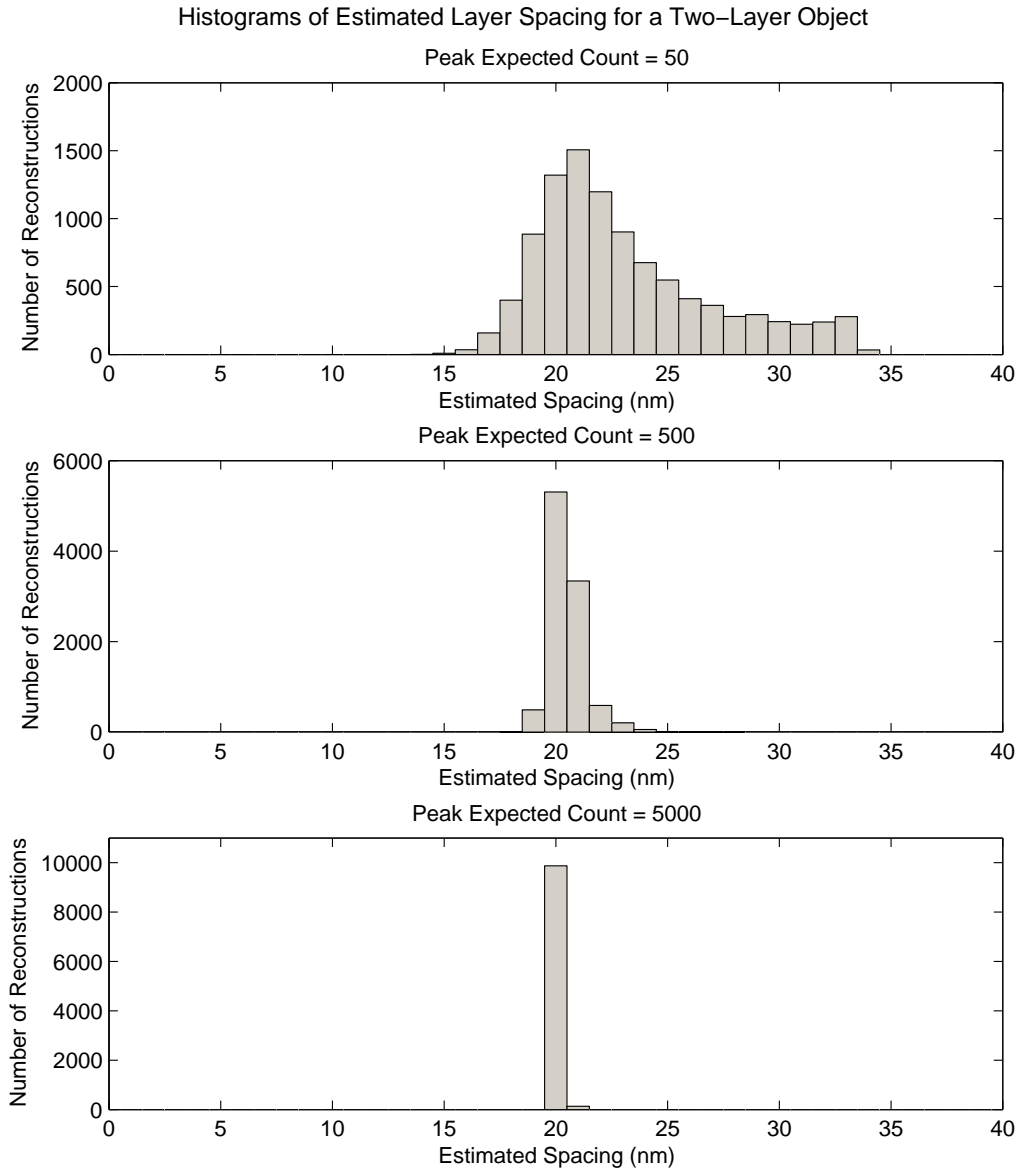
Figure 3·13: Histograms of the estimated distance between two layers located at $z = 100$nm and $z = 120$nm. The simulations are for peak expected signal levels of 50, 500 and 5000 photons. In each case 10000 simulations were run, each with a different realization of the Poisson noise. A spacing estimate of 20nm is clearly ideal for this object.

Two–Dimensional Histograms of Reconstruction Positions and Densities for a Two–Layer Object
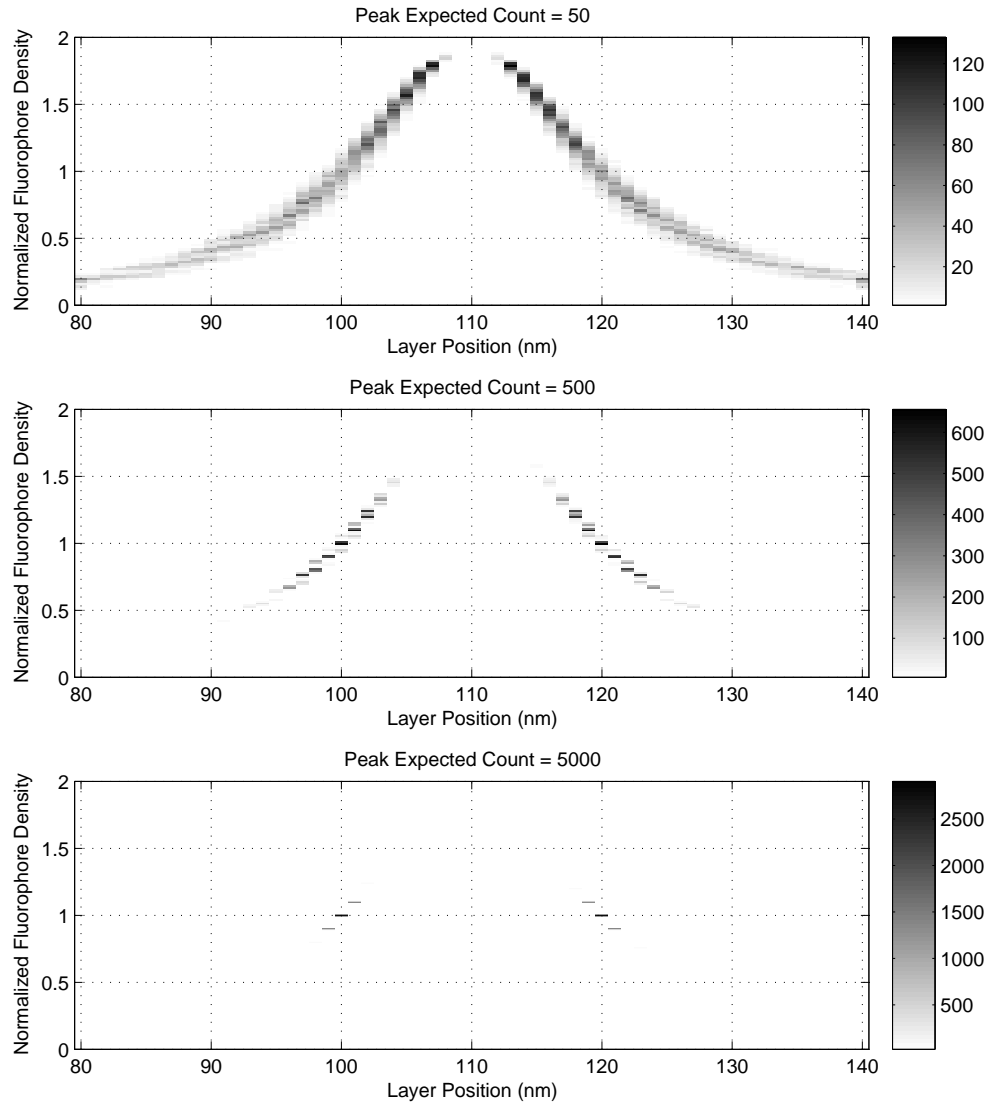
Figure 3·14: Two-dimensional histograms of the estimated layer positions and densities for a two-layer object. The layers are located at $z = 100$nm and $z = 120$nm. Simulations were run for peak expected photon counts of 50, 500 and 5000 photons. In each case 10000 simulations were run giving 20000 position-density data points.

Again a relatively low variability is seen in the 5000-count data, while high variability is seen in the 50-count data. The 5000-count data is close to the ideal case of two points at $(100, 1)$ and $(120, 1)$. The 50-count plot shows that erroneously high density estimates tend be positioned towards the midpoint of the two layers. This corresponds to the case where the two layers are estimated as a single strong intermediate layer with a low-density layer elsewhere. A high-density layer is not placed outside the two layers. The position correlation of Figure 3·12 shows that the low-density-layer estimate outside the true positions tends to be paired with a high-density-layer estimate between the true positions — two position estimates between the layers don't occur together just as two position estimates outside the layers don't occur together. The plots of Figure 3·12 and Figure 3·14 show that at high signal levels SSFM combined with the limited-component inversion approach can generally resolve layers spaced at 20nm. As the signal level decreases the two layers tend to be reconstructed as a single intermediate layer with a lower-density outlying layer.

3.4 Envelope Estimation

Up until this point it has been assumed that the fluorophore-emission spectral envelope (contained in the term $\mathcal{E}$) was known. This section gives a method for estimating the object when the spectral envelope is unknown.

3.4.1 Jointly Estimating the Object and the Spectral Envelope

When the fluorophore-emission spectral envelope is unknown it will also be estimated from the data. Thus the object and the spectral envelope will be jointly estimated. This will again be cast as a minimization problem and can be written as shown below — the operator $\mathcal{C}$ has been separated into $\mathcal{E}C$ and a general regularization cost has been included.

$$\hat{\mathbf{o}}, \hat{\boldsymbol{\epsilon}} = \operatorname*{argmin}_{\mathbf{o}, \boldsymbol{\epsilon}} \left\| [\Lambda(\mathbf{d})]^{-0.5}[\mathbf{d} - \mathcal{E}C\mathbf{o}] \right\|^2 + \alpha^2 \Psi_{\text{prior}}(\mathbf{o}) \tag{3.27}$$

Where $\mathcal{E}$ depends on $\boldsymbol{\epsilon}$ as shown in Equation 2.20 and $\Psi_{\text{prior}}(\mathbf{o})$ is defined by prior knowledge of the object. The first term in the minimization is the approximation to ML estimation

defined by Equation 3.9 and Equation 3.11. Again, a more sophisticated minimization procedure would be required if this approximation was not used. Such a procedure would give an exact, statistically-based solution to the joint estimation problem.

The difference between the observed data and the expected data will be written in its non-augmented form i.e. as a sum over the excitation patterns. This allows the spectral envelope vector ϵ to appear directly in the equation rather than embedded in $\mathcal{E}$.

$$\hat{\mathbf{o}}, \hat{\boldsymbol{\epsilon}} = \operatorname*{argmin}_{\mathbf{o}, \boldsymbol{\epsilon}} \sum_{j=1}^{J} \left\| [\Lambda(\mathbf{d}^{(j)})]^{-0.5} [\mathbf{d}^{(j)} - \operatorname{diag}(\boldsymbol{\epsilon}) C^{(j)} \mathbf{o}] \right\|^2 + \alpha^2 \Psi_{\text{prior}}(\mathbf{o}) \qquad (3.28)$$

The term $\Lambda(\mathbf{d}^{(j)})$ is similar to $\Lambda(\mathbf{d})$ (from Equation 3.9) except that it is calculated for the jth data set only.

Note that the observation equation is bilinear with respect to the two variables $\mathbf{o}$ and ϵ — i.e. the data is linearly related to the envelope if the object is held fixed and it is also linearly related to the object if the envelope is held fixed. This suggests that the problem can be solved using the coordinate descent algorithm (Bertsekas, 1999). This method involves successively minimizing over different subsets of coordinates in the cost function's domain. In this case the cost will be minimized over the object coordinates $\mathbf{o}$, followed by over the ϵ coordinates and back to the $\mathbf{o}$ coordinates, etc.

Minimizing over $\mathbf{o}$ while holding the envelope constant gives the following equation.

$$\hat{\mathbf{o}} = \operatorname*{argmin}_{\mathbf{o}} \left\| [\Lambda(\mathbf{d})]^{-0.5} [\mathbf{d} - \mathcal{E} C \mathbf{o}] \right\|^2 + \alpha^2 \Psi_{\text{prior}}(\mathbf{o}) \qquad (3.29)$$

This is the same problem addressed earlier in the chapter as $\mathcal{E}$ is held fixed. The same options are available for the prior cost, e.g. it may give Tikhonov regularization, be used to enforce non-negativity and/or include the p-norm to encourage sparsity.

The other part of the coordinate descent algorithm is to estimate the envelope ϵ while holding the object constant.

$$\hat{\boldsymbol{\epsilon}} = \operatorname*{argmin}_{\boldsymbol{\epsilon}} \sum_{j=1}^{J} \left\| [\Lambda(\mathbf{d}^{(j)})]^{-0.5} [\mathbf{d}^{(j)} - \operatorname{diag}(\boldsymbol{\epsilon}) C^{(j)} \mathbf{o}] \right\|^2 + \alpha^2 \Psi_{\text{prior}}(\mathbf{o}) \qquad (3.30)$$

This can be rewritten so that $\boldsymbol{\epsilon}$ appears simply as a vector. The new formulation uses the following diagonal matrix

$$\Upsilon^{(j)}(\mathbf{o}) = \text{diag}(C^{(j)}\mathbf{o}) \tag{3.31}$$

The minimization with respect to $\boldsymbol{\epsilon}$ can now be written in the familiar form below.

$$\hat{\boldsymbol{\epsilon}} = \underset{\boldsymbol{\epsilon}}{\text{argmin}} \sum_{j=1}^{J} \left\| [\Lambda(\mathbf{d}^{(j)})]^{-0.5} [\mathbf{d}^{(j)} - \Upsilon^{(j)}(\mathbf{o})\boldsymbol{\epsilon}] \right\|^2 + \alpha^2 \Psi_{\text{prior}}(\mathbf{o}) \tag{3.32}$$

The envelope appears modulated by the expected data. The cost function is a standard weighted LS term and a prior term that depends only on the object and thus does not affect the envelope estimate.

If the estimated envelope is allowed to take on an arbitrary form significant problems may arise. For example, the invertibility of $\Upsilon^{(j)}(\mathbf{o})$ means that in the one-measurement ($J = 1$) case, the square error term can be set to zero with $\boldsymbol{\epsilon} = \Upsilon^{-1}(\mathbf{o})\,\mathbf{d}$. In this case all the noise present on the data would appear in the estimate of the envelope. In order to obtain a smooth estimate of the envelope, it will be parameterized by a spline function. A spline function can be represented in the form shown below.

$$\boldsymbol{\epsilon} = \sum_{\ell=1}^{\mathcal{L}} (\boldsymbol{\varepsilon})_\ell \, \vartheta^{(\ell)}(\boldsymbol{\nu}) \tag{3.33}$$

Here $\vartheta^{(\ell)}(\boldsymbol{\nu})$ is the ℓth of $\mathcal{L}$ basis functions of the B-spline (De Boor, 2001). The vector $\boldsymbol{\varepsilon}$ contains the coefficients of the basis functions and is thus a parameterized representation of the envelope. Forcing the envelope to be a spline corresponds to having strong prior knowledge of its shape. Enforcing this knowledge solves any noise-stability and uniqueness problems that may have occurred when estimating the envelope.

The spline representation of the envelope can be easily incorporated into the reconstruction algorithm. A $M \times \mathcal{L}$ matrix Ξ can be defined where the ℓth column is the basis function $\vartheta^{(\ell)}(\boldsymbol{\nu})$. This allows Equation 3.33 to be written as a matrix equation.

$$\boldsymbol{\epsilon} = \Xi\boldsymbol{\varepsilon} \tag{3.34}$$

Now Equation 3.32 can be written in such a way as to estimate the spline parameters directly.

$$\hat{\boldsymbol{\varepsilon}} = \operatorname*{argmin}_{\boldsymbol{\varepsilon}} \sum_{j=1}^{J} \left\| [\Lambda(\mathbf{d}^{(j)})]^{-0.5} [\mathbf{d}^{(j)} - \Upsilon^{(j)}(\mathbf{o}) \, \Xi \, \boldsymbol{\varepsilon}] \right\|^2 \tag{3.35}$$

Again, this is simply a LS minimization problem and easily solved. Thus the coordinate descent algorithm used to estimate the envelope is defined by the two alternating estimation procedures of Equation 3.29 and Equation 3.35. All that remains is to define a starting point for the iteration.

Looking at Figure 3·2, it can be seen that the shape of the spectral envelope is clearly visible in the collected data. For this reason an initial estimate of the envelope will be found directly from the data and used to initialize the envelope-object estimation procedure. SSFM relies on interference effects producing oscillations in the measured spectra. Since the envelope can be expected to be uni-modal it is reasonable to say that the envelope and the interference-induced oscillations have significantly different supports for their Fourier transforms. This suggests the following method for getting an approximate estimate of the envelope.

1. Take the logarithm of each measured spectrum to convert the multiplicative relation between the envelope and the SSFM spectral oscillations to an additive relation.

2. Low-pass filter the results to remove the spectral-interference oscillations and any high frequency components of the noise.

3. Take the exponential to recover estimates of the envelope from each spectrum.

4. Normalize each of the envelope estimates to have a norm of 1.

5. Average the envelope estimates.

6. Fit a spline to this average.

The procedure outlined here is an application of homomorphic signal processing (Oppenheim and Schafer, 1989). The critical assumption in this approach is that the logarithm
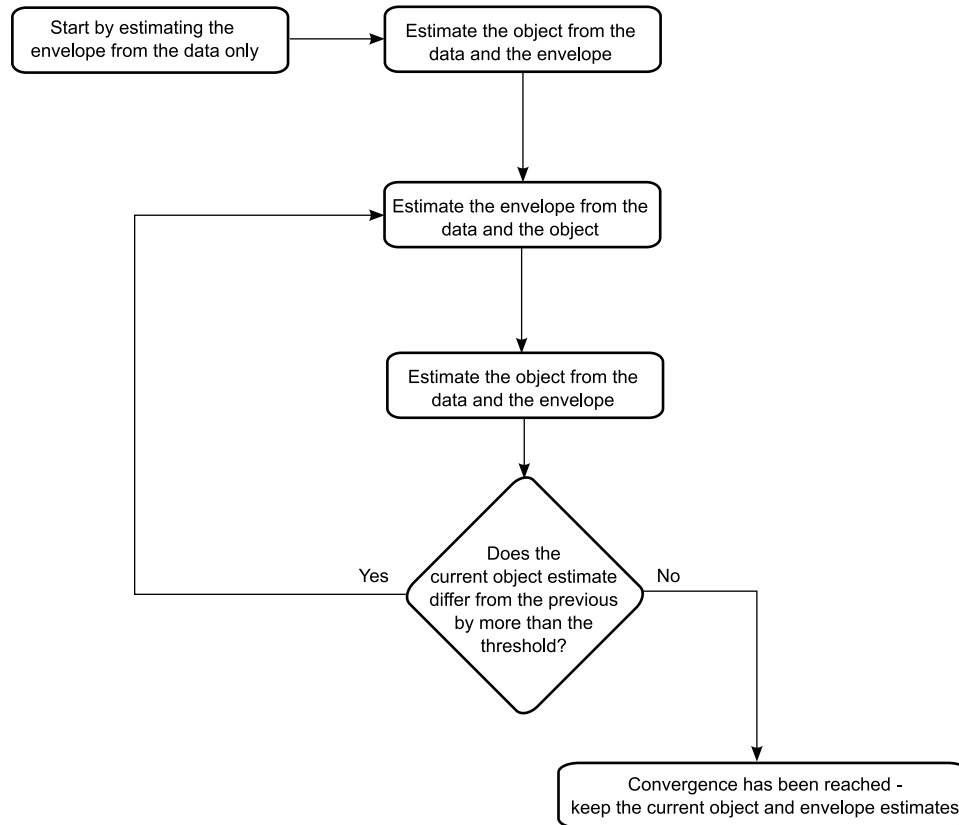
Figure 3·15: Flow diagram showing the coordinate descent algorithm used to jointly estimate the fluorophore-emission spectral envelope and the object.

of the Fourier transform of the envelope and the logarithm of the Fourier transform of the spectral-oscillation data have distinct supports so they can be separated using a filter. The simulation results shown later indicate that the method works quite well.

Once this initial estimate of the envelope is found, the coordinate descent algorithm can be started by first estimating the object (Equation 3.29), then the envelope (Equation 3.35), etc. This iterative algorithm is illustrated in Figure 3·15. Each time the envelope is estimated it will be scaled to have a norm of 1 as there is a scale ambiguity between the object and the estimate. That is, if the object is scaled up by some constant and the envelope down by the same amount, the solution will be equally valid.
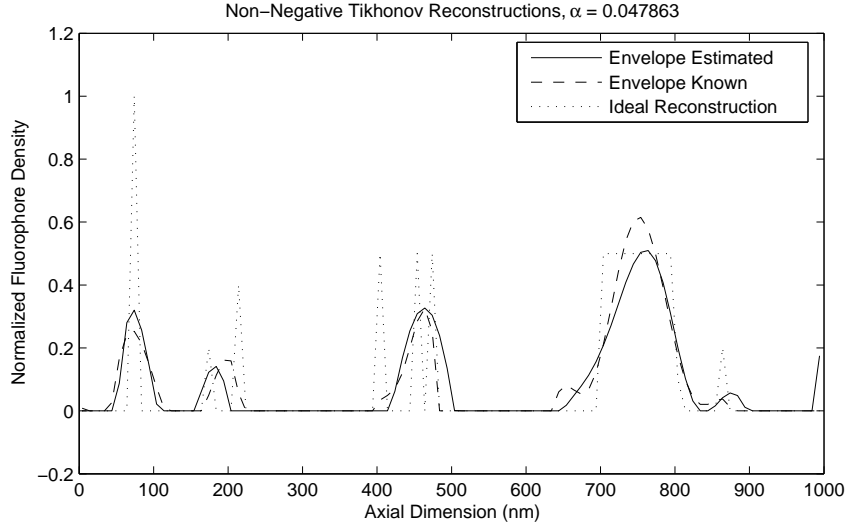
Figure 3·16: Non-negative Tikhonov reconstructions of the test object from the simulated data. In one case the envelope is assumed known and in the other it is estimated jointly with the object. The ideal reconstruction is also plotted.

3.4.2 Simulation Results

To illustrate the effectiveness of the envelope estimation method, simulations will be performed. Up until this point all the simulations in this chapter have assumed that the envelope is known. These known-envelope reconstructions can be used as one part of the iterative algorithm described to jointly estimate the object and envelope, the other part involving fitting a spline to estimate the envelope. A cubic spline was used to fit the envelope. This spline had twelve uniformly spaced knots, six of which were used to uniformly partition the spectral axis into five intervals (six of the intervals fall outside the spectral range considered). The result is $\mathcal{L} = 8$ spline basis functions over the spectral axis. It should be noted that the envelopes will be estimated from noisy data and that the form of the true envelope (which is the same for all simulations) is a Gaussian plus a linear term, not a spline function.

The best non-negative Tikhonov reconstruction of Figure 3·7 can be repeated where the envelope is also estimated. The result is shown in Figure 3·16. Care was taken with

the envelope scaling to assure that the regularization parameter α had a consistent effect between the two simulations (i.e. the envelope was normalized to the same level in each case). The coordinate-descent algorithm was assumed to have converged when relative energy change between iterations was less than 1%. It can be seen that estimating the envelope does change the object estimate but only slightly. Quantitatively, the estimated-envelope reconstruction's square error was only 2.3% higher than the square error in the known-envelope case.

A two-layer reconstruction was also performed for a peak expected count of 500 and where the envelope was unknown. The object has layer positions of 100nm and 120nm with densities of 1 in both layers. The coordinate-descent algorithm was assumed to have converged when successive position estimates and successive density estimates both varied by less than 1%. The data used produced position estimates of 102nm and 122nm and density estimates of 1.20 and 0.80 when the envelope was known. When the envelope had to be estimated the positions estimated were 102nm and 122nm (all positions were estimated to the nearest nanometer) and the densities estimated were 1.19 and 0.79. This also shows only a slight change when the envelope was unknown. Thus for the illustrative simulations run here, the effect of an unknown envelope was mild.

The estimates of the envelope are shown in Figure 3·17. One plot shows the initial estimates of the envelope found by applying the filtering operation described earlier to the raw data. The final estimate of the envelope is also shown for the two simulations and is seen to differ negligibly from the true envelope. The percentage difference between the true envelope and the estimated envelopes is plotted and shown not to exceed 3%.

3.5   SSFM Compared to Other Microscopical Imaging Systems

To this point the simulation results have been for SSFM systems. The results have shown a single-layer localization precision of better than 1nm, an ability to separate two layers spaced 20nm apart (given an adequate signal level) and resolution of an unconstrained object on the order of 100nm. To determine the utility of SSFM it is critical to know how
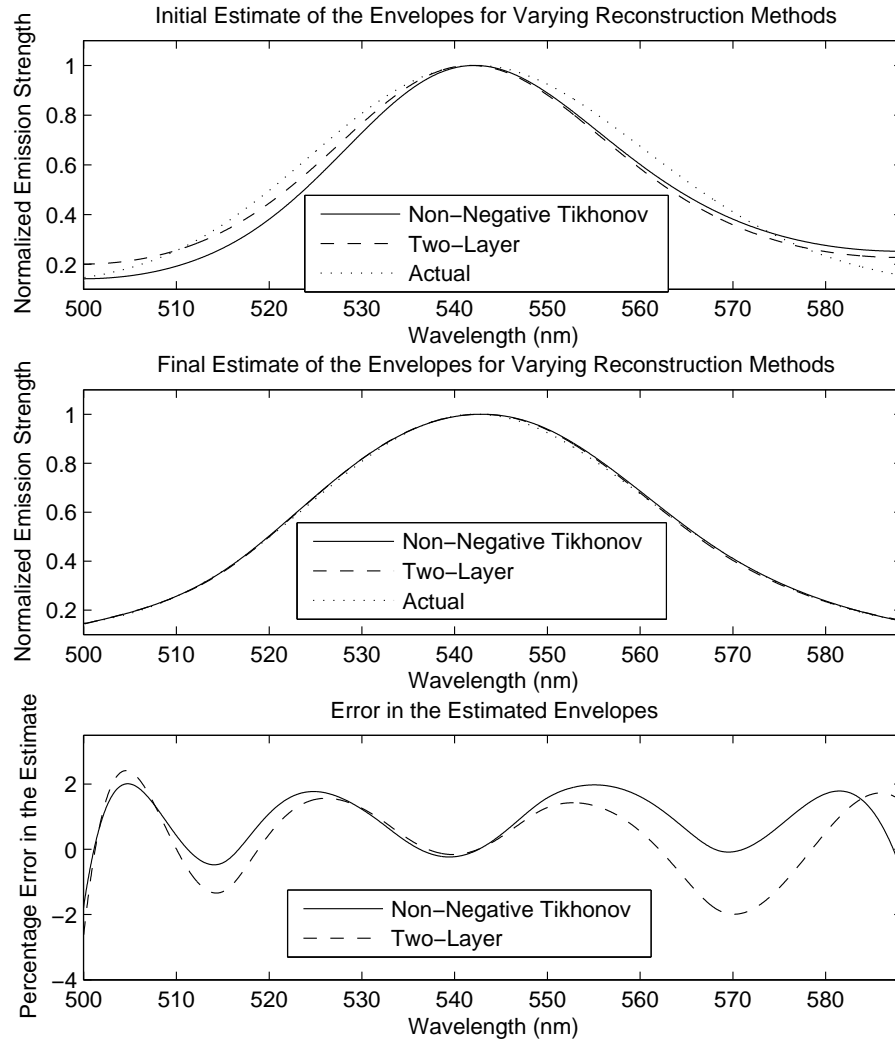
Figure 3·17: Estimated envelopes for the two simulations considered —
a two-layer reconstruction case and a non-negative Tikhonov reconstruc-
tion case. The initial estimates of the envelope, from the data, are shown
along with the true envelope. The final estimates of the envelope (given
after the iterative envelope-object estimation algorithm has converged) are
shown along with the true envelope. The percentage error between the final
estimates and the true envelope is also shown.

this compares to other available microscopy systems. This section addresses that question.

3.5.1 Similar Reconstruction Methods Applied to Other Systems

In this section some of the reconstruction methods defined above are applied to other high resolution microscopy systems — specifically confocal microscopy, I^5M (Incoherent Interference Illumination, Image Interference Microscopy) and 4Pi type C microscopy. As above, the instruments will be applied to a purely axial object. The results contained in this section are also found in the paper (Davis et al., 2004b).

The three microscopy methods investigated here are all linear and shift-invariant which means that each can be characterized by its point spread function. The methods for calculating these are well known (van der Voort and Brakenhoff, 1990; Hell and Stelzer, 1992a; Nagorni and Hell, 2001b). The axial PSFs (calculated by integrating over the lateral plane) for the three systems were used to generate data for a two-layer object and are shown in Figure 3·18. All simulated data and reconstructions were calculated on a 5nm grid. The instrument parameters used to generate the PSFs were chosen to be consistent with practically optimized biological imaging. Due to the different operating principles of SSFM, these parameters may not always be directly comparable (e.g. SSFM can achieve high axial resolution with low numerical apertures while a confocal system needs a high NA to optimize its resolving capabilities). In all cases the maximum expected photon count was taken to be 150. The signal level is also difficult to compare to SSFM since the SSFM system described does not scan and can therefore be expected to collect higher-count data in the same image acquisition period.

The spacing between the two object layers was varied. For 200nm spacing, non-negative Tikhonov and two-layer reconstruction (as in Section 3.3) results are shown in Figure 3·19. It can be seen that layers at this spacing are resolved in the non-negative Tikhonov reconstructions (although they are broadened) and in the two-layer reconstructions. The reconstructions are not as good in the confocal case. This is to be expected as the I^5M and 4Pi systems are specifically designed to produce higher axial resolution.
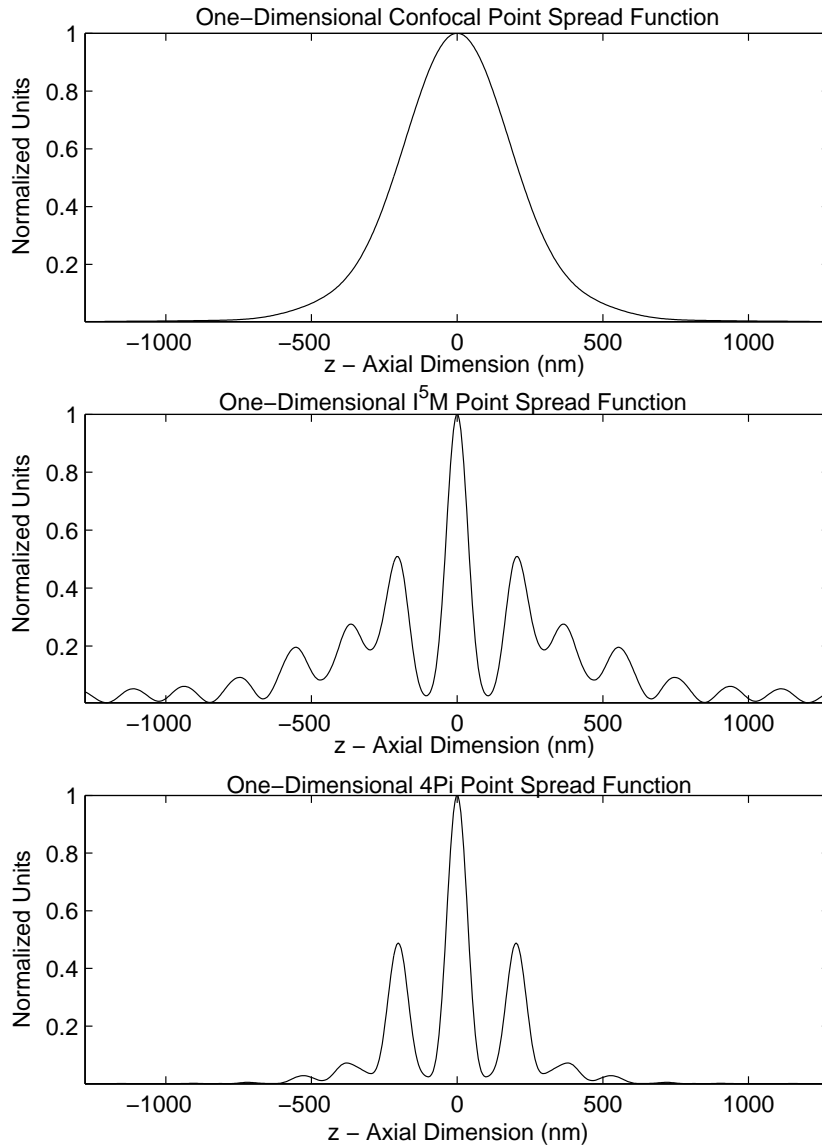
Figure 3·18: One-dimensional point spread functions for the example confocal, I^5M and 4Pi systems. In all systems the excitation wavelength is 488nm, the detection wavelength is 500nm, the index of refraction n is 1.51 and the numerical aperture is 1.31
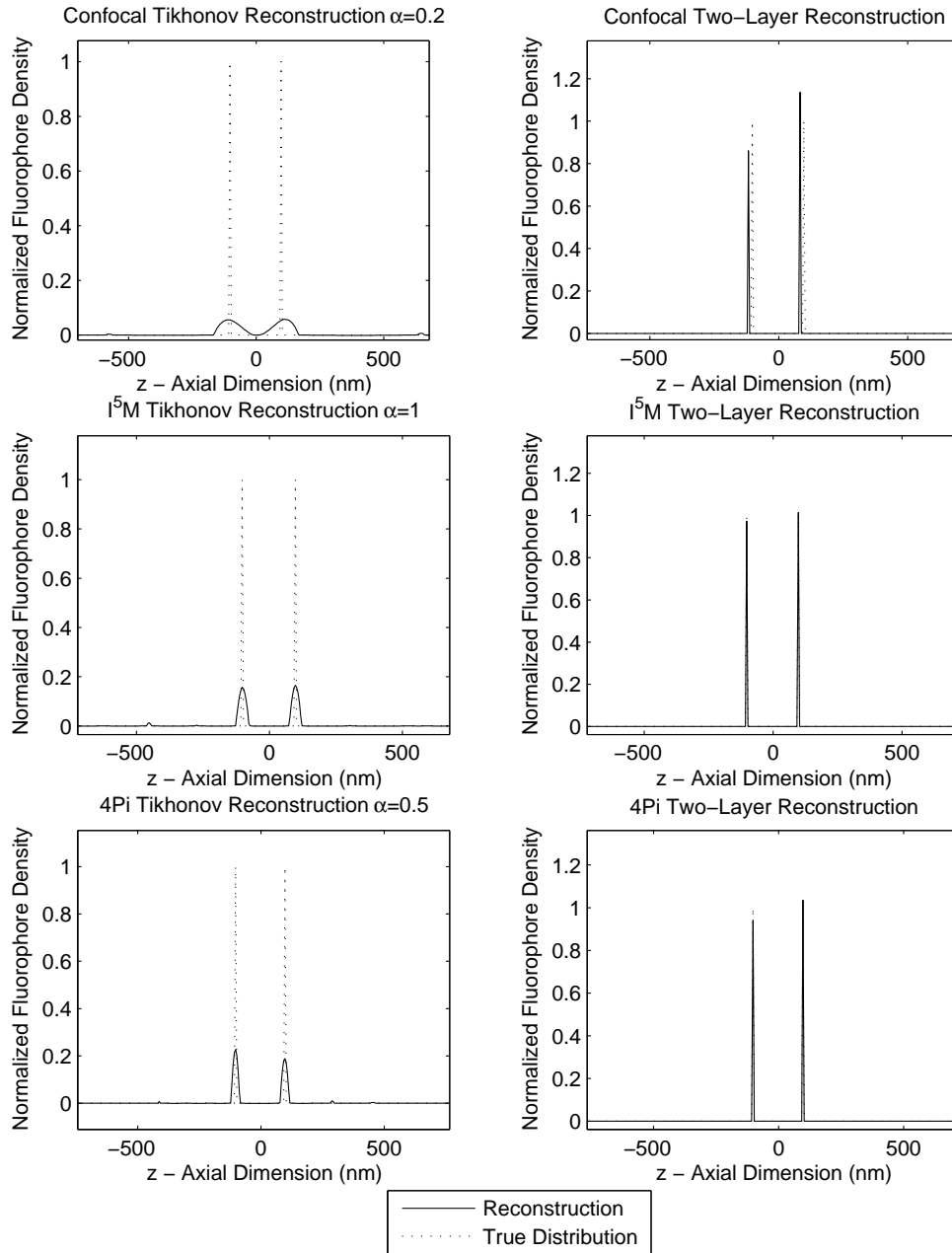
Figure 3·19: Non-negative Tikhonov and two-layer reconstructions for the example confocal, I^{5}M and 4Pi systems where the two layers are separated by 200nm. The Tikhonov regularization parameters were chosen manually and are displayed on the graphs.

The layer spacing was reduced to 20nm and the resulting reconstructions are shown in Figure 3·20. The advantages of limited-component reconstruction are apparent as the layers are not resolved in the non-negative Tikhonov cases but they are in the 4Pi and I^{5}M two-layer reconstructions. More importantly, these simulations demonstrate 4Pi and I^{5}M systems perform in a similar manner to SSFM in limited-components reconstructions. It is also well known that standard microscopy systems can localize single particles to nanometer precision (Schmidt et al., 2000; Albrecht et al., 2002; Ober et al., 2004; Ram et al., 2005; Aguet et al., 2005; Ram et al., 2006).

3.5.2 Fourier Domain Analysis

A Fourier analysis can be used to give rough limitations on the resolution of a SSFM system. A very simple model of the system will be used but it nevertheless captures the general behavior of a SSFM system. Standing wave excitation will be assumed.

$$|h_e(z)|^2 = 1 + |\rho_e|^2 + 2\mathrm{Re}\{\rho_e\}\cos\left(\frac{4\pi n}{\lambda_e}z + \phi_e\right) \tag{3.36}$$

This is the model of Equation 3.1

Similarly, the detection pattern is determined by interference between the two paths. The optical path difference is $2z$ — twice the distance from the fluorophore to the mirror. With respect to Equation 2.8, $\tau(z) = \frac{2nz}{c}$ and $\nu = \frac{2\pi c}{\lambda_d}$ where c is the free-space speed of light and λ_d is the detection wavelength.

$$|h_d(z;\lambda_d)|^2 = 1 + |\rho_d|^2 + 2\mathrm{Re}\{\rho_d\}\cos\left(\frac{4\pi n}{\lambda_d}z + \phi_d\right) \tag{3.37}$$

The observation equation involves projecting the object on to $|h_e(z)|^2|h_d(z;\lambda_d)|^2$. Taking the excitation-detection product results in sinusoids with frequencies that are the sums and differences of the factors' frequencies. This is shown in Figure 3·21. The maximum frequency observed is $4\pi n(\frac{1}{\lambda_e} + \frac{1}{\lambda_d})$. As will be seen in Chapter 5, this is the same cutoff as a 4Pi system.

The simple model applied here is readily represented in the Fourier domain as it is
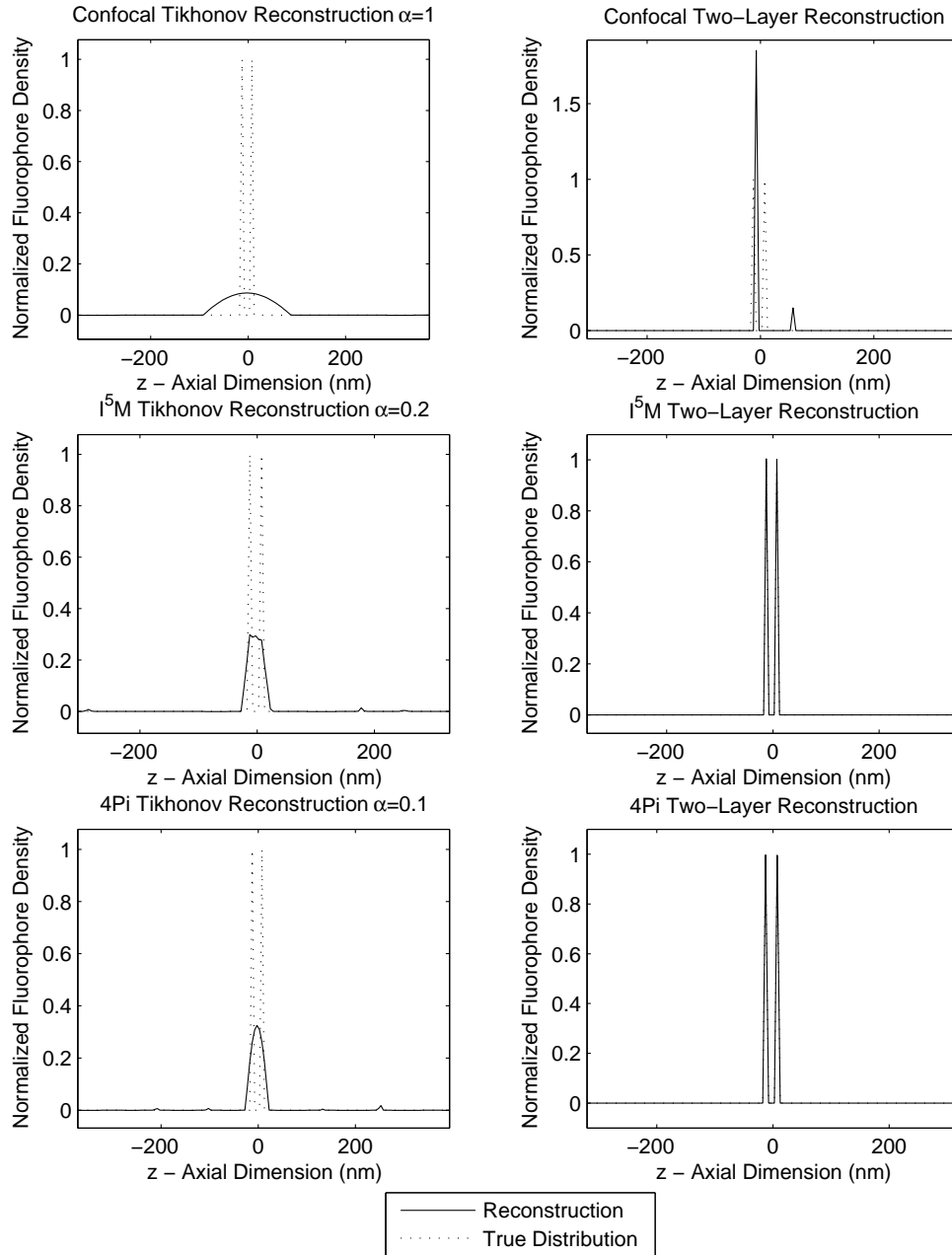
Figure 3·20: Non-Negative Tikhonov and two-layer reconstructions for the example confocal, I^{5}M and 4Pi systems where the two layers are separated by 20nm. The Tikhonov regularization parameters were chosen manually and are displayed on the graphs.
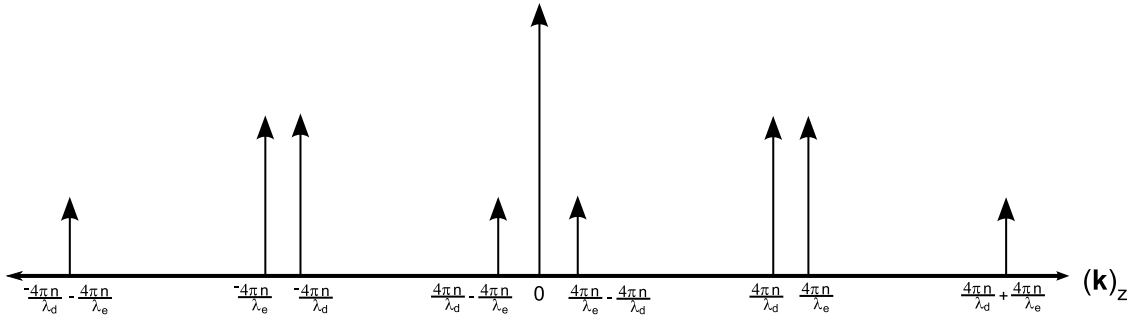
Figure 3·21: Fourier-domain representation of the observable spatial frequencies in the simplified one-dimensional SSFM model. The diagram is not to scale and the impulse amplitudes give only a qualitative indication of the relative observation strengths. Phase relations between the observed frequencies are not shown either.

comprised of trigonometric functions. A more accurate model must include such things as focusing effects and as a result will not be so simply represented in Fourier space. The next chapter shows how such a model can be constructed and in Chapter 5 it will be seen that this basic Fourier analysis still holds for a comprehensive model.

3.6   Evanescently-Excited SSFM

This section introduces a recently-proposed modification to SSFM. This work is still in its preliminary stages and the results presented here should be interpreted in that light. Dr. P. Scott Carney of the University of Illinois at Urbana-Champaign is acknowledged for his investigation into the theory of this system — the figures generated in this section use his physical models and simulation code.

3.6.1   Using Evanescent Excitation in SSFM Systems

Evanescent waves are used in Total Internal Reflection Fluorescence (TIRF) microscopy (Axelrod et al., 1982) to ensure that only a very thin slice of the object is illuminated. Fluorescent emission does not occur outside this illumination volume and thus only a very thin plane is imaged within the object. Evanescent waves can be produced at the boundary of two media with different optical properties. In TIRF microscopy they are produced by illuminating the underside of this boundary with light incident at an angle beyond the
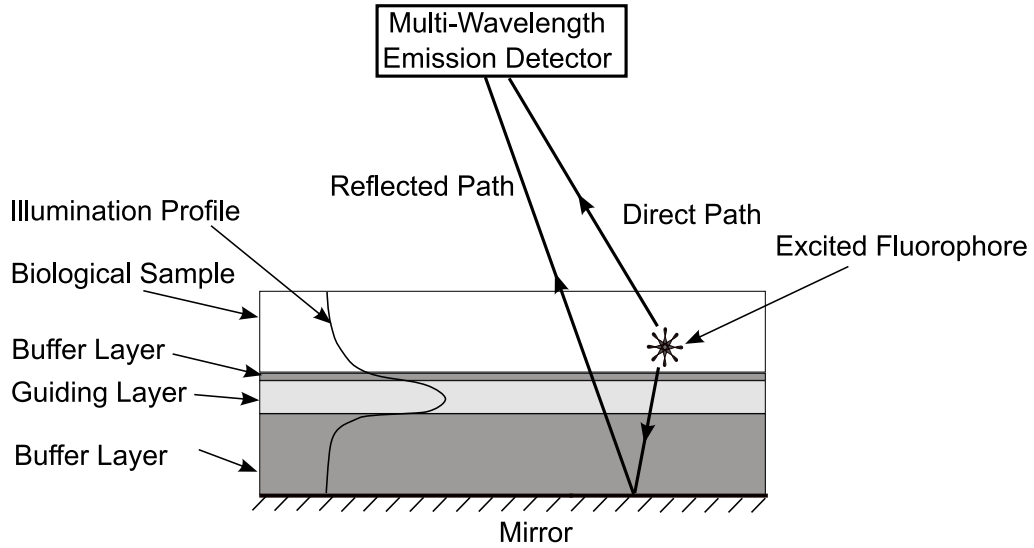
Figure 3·22: An illustration of evanescently-excited SSFM. The object is placed on top of a waveguide structure and the evanescent field produced by the guided light excites fluorophores near the bottom of the object preferentially. Detection is performed as in standard SSFM.

critical angle. Evanescent waves may also be produced and controlled in a waveguide structure as discussed in (Worth et al., 2001). In either case, the intensity of the field decays exponentially away from the boundary. If a laterally oriented boundary is assumed to be at $z = 0$ then the evanescent exciting intensity can be described as follows.

$$|h_e(z)|^2 \propto e^{-2\frac{2\pi n}{\lambda_e}\varsigma z} \tag{3.38}$$

Here the constant ς determines the rate of decay of the field and the first 2 in the exponent comes from squaring the field to get the intensity. The decay rate ς is itself determined by the design of the system — in TIRF microscopy it depends on the incidence angle of the light at the boundary and in waveguide structures it depends on the waveguide mode used. Evanescent excitation can be incorporated into a SSFM system as illustrated in Figure 3·22.

As shown in Figure 3·22, the evanescent field from the guided light penetrates into the object causing fluorophores near the object-waveguide boundary to be excited. The emissions from these fluorophores are imaged using the standard SSFM methodology. It

should be noted that a smaller volume is imaged in evanescently-excited SSFM due to the decaying excitation.

The excitation pattern can be altered in standard SSFM by changing the phase on the standing wave, as discussed in Section 3.1. The evanescent excitation can also be altered to gain additional information about the object. At a minimum, the object can be illuminated with a uniform field in addition to the evanescent field. If a multi-mode waveguide is used, each mode will generally produce a different evanescent field. Changing the mode(s) propagating in the waveguide thus allows the excitation field to be varied.

Some early simulation results for evanescently-excited SSFM will now be presented. The model was developed by Dr. P. Scott Carney and consists of an excitation modeled by Equation 3.38 and a detection model similar to Equation 3.37. Beyond these simulations there will be little more analysis of evanescently-excited SSFM presented, as this technology is still in the preliminary stages of development.

3.6.2 Preliminary Simulation Results

The simulations presented here consider a two-layer object with the layers positioned at $z = 37$nm and $z = 163$nm. The excitation light is at $\lambda_e = 500$nm and the detection wavelength ranges from 470nm to 525nm. The fluorophore emission is assumed to be uniform over this range. The excitation field has a decay rate of $\varsigma = 0.5$ which means that the excitation intensity has dropped by a factor of 2 after approximately 80nm. The resulting data is shown in Figure 3·23.

The difference between evanescent excitation and uniform excitation can be clearly seen in Figure 3·23. With the uniform illumination both layers contribute equally to the data while with the evanescent illumination the excitation field is significantly higher at $z = 37$nm than at $z = 163$nm. This results is a significantly lower contribution from the $z = 163$nm layer. This distinction of the two layers in the two sets of data indicates that using these multiple excitation patterns may be very valuable in resolving the two layers.

The two-layer object will now be reconstructed from the 'Noisy Data' plots of Fig-
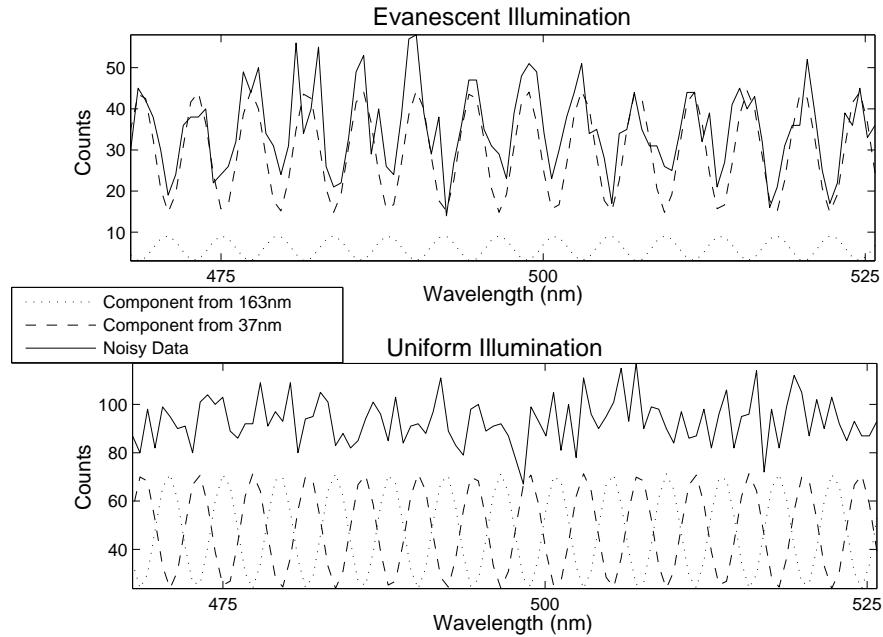
Figure 3·23: Simulated data for an evanescently-excited SSFM system imaging a two-layer object. The object has layers at $z = 37$nm and $z = 163$nm. Two excitation patterns are considered — a uniform excitation and an evanescent excitation with a decay rate of $\varsigma = 0.5$. In both cases the data that would result from each layer individually is plotted. The 'Noisy Data' is what would be actually observed and is the sum of the components from the two layers, with Poisson noise included.

ure 3·23. As can be seen from the figure, this is not high-signal data. The object is reconstructed on the range $z = 0 - 250$nm with 200 sample points within this range. A general reconstruction method will first be considered. The Truncated-Singular-Value Decomposition (TSVD) (Hansen, 1998) is an inverse of the operator $\mathcal{C}$ but with weakly passed components set to zero. Thus only components of the object that are passed with a gain above a specified threshold are estimated in a TSVD reconstruction. This process allows stability in the presence of noise and the threshold plays the role of the regularization parameter α. A TSVD reconstruction of the two-layer object is shown in Figure 3·24. The threshold was set to a level that was consistent with the signal level. Any negative values in the reconstruction were set to zero as the fluorophore density is known to be non-negative.

The reconstruction of Figure 3·24 shows that the two layers are clearly resolved. Each
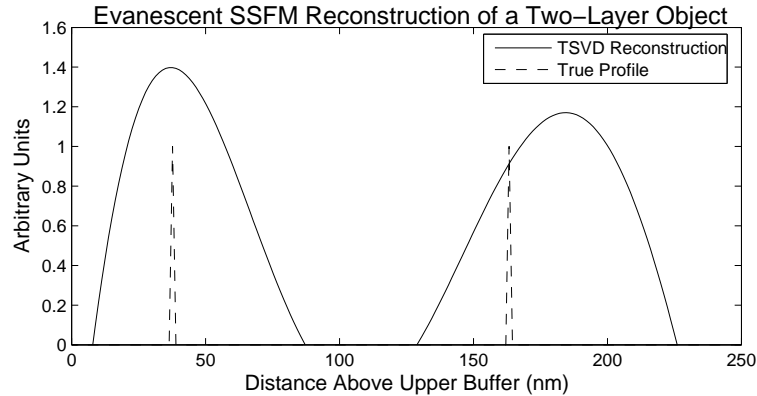
Figure 3·24: Truncated-singular-value-decomposition reconstruction of the two-layer object from the evanescently-excited SSFM data of Figure 3·23. Any negative values in the reconstruction were set to zero. The object is also plotted but scaled down so that the reconstruction is visible on the same scale.

is spread over an axial range of less than 100nm. Since the TSVD is a general reconstruction method, this gives a reasonable idea of the resolution that can be expected using evanescently-excited SSFM.

Limited-component reconstruction methods were applied to SSFM in Section 3.3 — they can also be applied to the evanescently-excited SSFM problem. A single-layer object, with the fluorescent layer positioned at $z = 37$nm, will first be considered. The signal level is the same as that displayed in Figure 3·23 and following the approach used earlier, many reconstructions (with different realizations of the noise) will be performed. A histogram of the estimated layer positions is shown in Figure 3·25.

The histogram in Figure 3·25 indicates that precision on the order of a nanometer is again achieved when estimating the position of a single layer. The next object considered will consist of two layers positioned at $z = 37$nm and $z = 81$nm. Once again many simulations will be run in order to investigate the statistics of the reconstruction. Rather than plotting a two-dimensional histogram of the position estimates (as in Figure 3·12), Figure 3·26 has the two position estimates plotted on the same axis. While this does mean that the correlation between the two position estimates will no longer be visible, the histogram will give a graphical representation of the spread of the two position estimates.
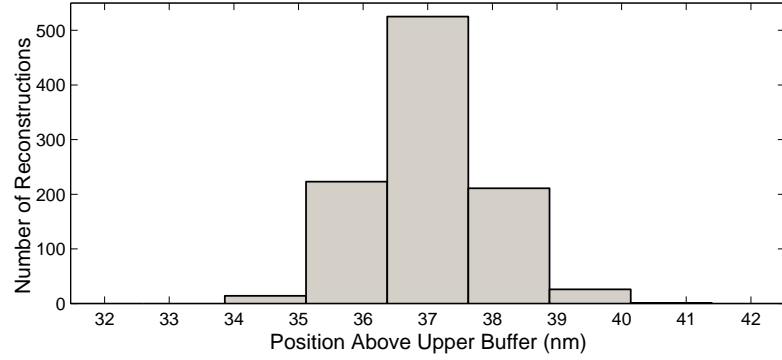
Figure 3·25: Histogram of the estimated layer positions for an evanescently-excited SSFM system imaging a single-layer object. The layer is positioned at $z = 37$nm and 1000 simulations were run, each with a different realization of the Poisson noise. The position is estimated on a 1.25nm grid.
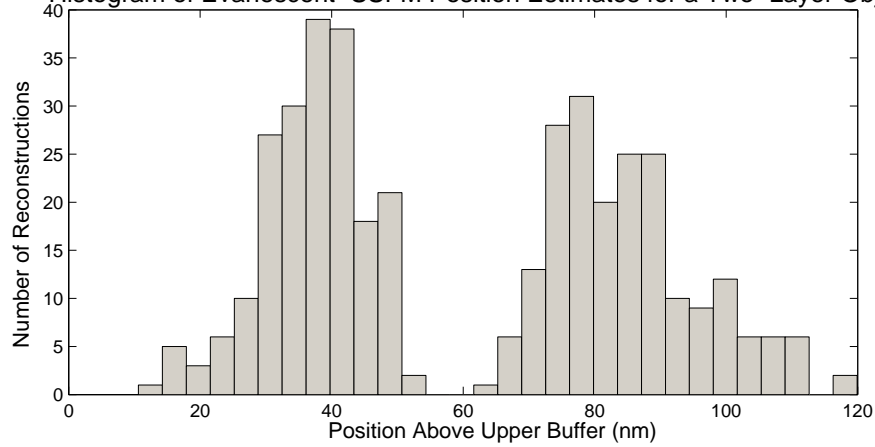
Figure 3·26: Histogram of the estimated layer positions for an evanescently-excited SSFM system imaging a two-layer object. The layers are positioned at $z = 37$nm and $z = 81$nm and 200 simulations were run, each with a different realization of the Poisson noise. The positions are estimated on a 1.25nm grid. The two position estimates are plotted on the same axis.

The histogram of Figure 3·26 shows two clearly separated layers. This indicates that these layers, which are spaced 44nm apart, can be distinguished using evanescently-excited SSFM. Each estimate varies by tens of nanometers.

Comparing the general estimation results to those presented in Section 3.2 indicates that evanescently-excited SSFM and standard SSFM have a similar resolution. Similarly, the preliminary limited-component reconstructions presented here indicate a similar performance to the standard-SSFM limited-component reconstructions shown in Section 3.3. Again, it should be noted that these results are preliminary and further investigation into evanescently-excited SSFM is required.

3.7 Summary

The approach taken in this chapter was to characterize the performance of SSFM through illustrative simulations. In Section 3.2 some standard reconstruction approaches were applied to simulated data from a test object. The resulting images indicated that the expected resolution of a SSFM system is not any higher than in a 4Pi system. There are other reconstruction approaches that have been applied in microscopy, e.g. POCS (Homem et al., 2002), modifications of the Richardson-Lucy algorithm (Conchello and McNally, 1996) or Maximum Entropy (ME) methods (Frieden, 1972; Verveer et al., 1999), which could also be applied here. Comparisons and surveys of deconvolution in microscopy also exist (van Kempen et al., 1997; Swedlow et al., 1997; Verveer et al., 1999; Wallace et al., 2001). Rather than attempting more examples with different deconvolution methods, the following chapters will characterize SSFM and show that it does not pass any more information than a modified 4Pi system.

In Section 3.3 a reconstruction method is developed that can be applied when the object is known to have only a limited number of components. It has been shown that these components can be estimated very precisely. This precision was what originally led to interest in SSFM as the spectral signature measured allows a fluorescent layer to be localized with nanometer accuracy. The accuracy expected when estimating two layers

was also explored here. While the results found with SSFM were impressive, they are similar to those achievable by 4Pi and other microscopy methods. This indicates that the high localization accuracy is produced by the prior knowledge of the object rather than any increased imaging capacity of the SSFM instrument. A SSFM system using an alternative evanescent excitation mechanism was also introduced.

The spectral emission profile of the fluorophores may not be precisely known and can introduce uncertainty in the model used in the reconstruction process. In Section 3.4, a method was developed to estimate this envelope accurately in conjunction with the object imaged. The method was developed specifically for this application. Similar general approaches, known as blind deconvolution, exist for image reconstruction when the system model is not known or is only partially known. These have been applied in microscopy (Holmes, 1992; Holmes and O'Connor, 2000) but are not investigated here as the envelope is estimated accurately with the method given.

In short, this chapter provides a strong indication that SSFM may not provide the capacity for improved resolution. SSFM may not be revolutionary in terms of image quality but it may be useful for certain applications as it is relatively easy to implement (especially when compared to other interferometric techniques). SSFM also provides imaging capability with only one to three measured spectra as compared to scanning systems which require a measurement for each image pixel. It should also be noted that the numerical experiments performed in this chapter are only examples that suggest the resolution limitations of SSFM — the origins of these limitations will be explored and quantified in subsequent chapters.

Chapter 4

Relation of SSFM to Known Imaging Systems

The previous chapter showed that despite promising initial results with single component objects, the SSFM system does not seem to provide a significant advantage over certain other forms of light microscopy. This chapter shows how these simulation results can be explained theoretically. This analysis relies heavily on optical transfer function analysis and recognizing the commonalities between SSFM and other forms of microscopy, in particular I^5M and 4Pi Type C microscopy.

In order to calculate OTFs, the point spread functions of several systems are found. This chapter explains, in detail, how to calculate fully-polarimetric point spread functions for these high-numerical aperture, far-field, light microscopy systems. While this process is well known to experts, a thorough and detailed explanation is generally omitted from papers published in light microscopy due to page limitations.

4.1 SSFM Compared to 4Pi, I^5M and Standing Wave Microscopy

A simple schematic of a 4Pi Type C microscope is shown in Figure 4·1. It can be seen that the excitation light arrives from both sides of the sample. This is similar to SSFM where the excitation light travels in both directions (the incident light and the reflected light). For both methods, the two directions of propagation are used to create interferometric structure in the excitation light. Similarly, the detection is from both sides of the sample in a 4Pi system. Again, this is similar to the SSFM system where emitted light can travel two paths to the receiver. These similarities provide a convenient framework for examining the SSFM system. In fact, the SSFM system can be regarded as a 4Pi Type C system but with several key differences.
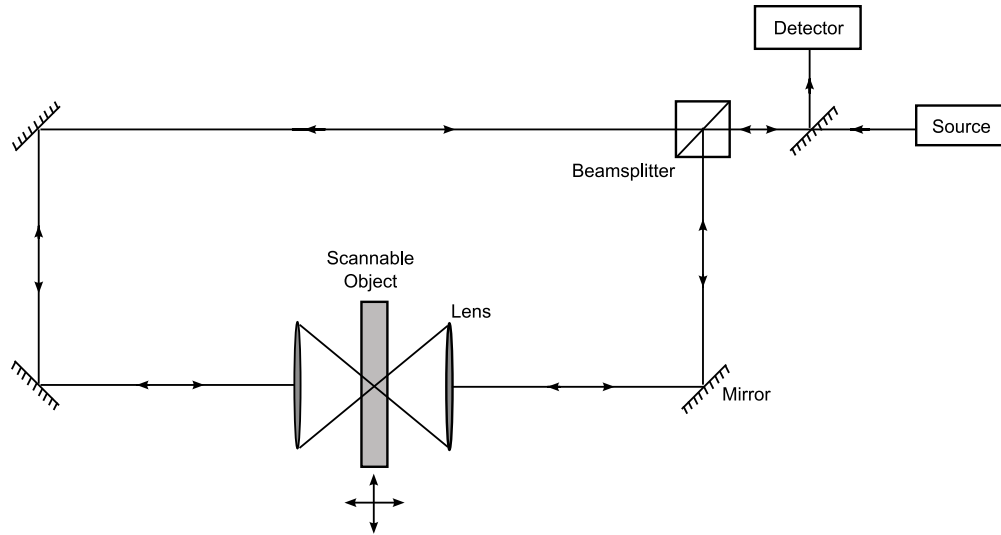
Figure 4·1: Basic schematic of a 4Pi Type C microscope. Excitation and detection are both through twin opposing objective lenses. The two paths to the sample will generally have the same length although this is not indicated in the diagram.

1. SSFM has a spectral detector while 4Pi instruments typically detect light at only a single wavelength.

2. SSFM has a path-length difference between the two routes from object to detector and/or from excitation source to object.

3. 4Pi systems involve spatial scanning while SSFM may not necessarily include axial scanning.

4. The two objective lenses in a 4Pi system are generally aligned so that their foci match. Due to the spacer above the mirror, the SSFM system cannot perfectly focus the direct and reflected light simultaneously.

Note that a 4Pi system typically only detects at a single wavelength, although two-color measurement has been demonstrated in order to measure an object stained with two different fluorophores (Kano et al., 2002). An I^5M system is very similar to a 4Pi system (and hence to SSFM) with the key difference being that the excitation uses incoherent, collimated light. The connection between SSFM and 4Pi will be examined in this section

but it should be understood that it is not a large leap to an I^5M system. 4Pi microscopy is also a generalization of standing wave fluorescence microscopy — SWFM is essentially a 4Pi instrument with zero-numerical-aperture lenses for the excitation and detection only from a single direction. The connection between a 4Pi system and a single objective over a mirror (as in SSFM) has also been noted in the literature (van de Nes et al., 2004).

This chapter details the calculations necessary to characterize a 4Pi Type C microscope with spectral detection, a path-length difference and possibly mismatched foci. This system is a generalized SSFM system. This spectral system can be treated as a series of single-wavelength systems operating in parallel.

4.2 Characterizing a Single Lens

This section contains a systematic analysis of how the diffraction pattern of a high-numerical-aperture lens, in the focal region, can be used to characterize the excitation and detection of fluorescent molecules. This analysis is fully polarimetric and includes details such as the effect of a finite detector size and the orientation of the lens. The analysis for more complex systems (such as SSFM) is based on these results.

4.2.1 Excitation Through a Lens

The excitation of fluorophores within the object is generally achieved using light focused through a lens. This section shows how that light's spatial distribution can be described and includes effects such as the polarization state of the exciting light and the direction of illumination. In the previous chapter a discrete notation was used for the purposes of computer simulation. The results shown here are continuous but it is understood that sampling can be used to create a discrete model as shown in Chapter 2. While scalar optics were assumed (for simplicity) in Chapter 2, a fully vectorial analysis will be used here as polarization effects can become significant in high NA systems. Complex notation will also be used to describe the fields calculated.

The electric field produced by the lens, in the region of the focal spot, is called the

Amplitude Spread Function (ASF) and will be denoted by $\vec{a}_x(\mathbf{r}; \lambda_e)$. The origin $\mathbf{r} = \mathbf{0}$ is defined as the geometric focal point. Details for calculating $\vec{a}_x(\mathbf{r}; \lambda_e)$ are given in (Wolf, 1959; Richards and Wolf, 1959). The subscript x indicates that the light incident on the lens is x-polarized and λ_e gives the excitation wavelength (note that the term ν has been used to define the spectral axis but the wavelength λ can also be used). The numerical aperture and the refractive index of the object also determine $\vec{a}_x(\mathbf{r}; \lambda_e)$ but they are generally fixed and will not be included as arguments of the function. Note that $\vec{a}_x(\mathbf{r}; \lambda_e)$ is a vector function and therefore has x, y and z components. All of the lenses described here will be assumed aplanatic (Born and Wolf, 1999) unless otherwise specified.

Illumination Polarization and the Excitation Pattern

To quantify the effects of various changes to the microscope system, some spatial transformations will be used. These operate on (x, y, z) vectors or equivalently $\mathbf{r}$ vectors — both of these notations will be used in this work. The first transformation used is a $90°$ rotation in the x-y plane. This rotation takes a vector lying along the x axis and rotates it so that it lies along the y axis.

$$T_y = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad T_y^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{4.1}$$

Thus if the lens is illuminated with y-polarized light, the following pattern is produced in the region of the focal spot.

$$\vec{a}_y(\mathbf{r}; \lambda_e) = T_y \, \vec{a}_x(T_y^{-1}\mathbf{r}; \lambda_e) \tag{4.2}$$

Notice that the transformation is applied the vector function to ensure the polarization state changes. The inverse is applied to the function's argument so that a point on the y axis, with y-polarized light, produces the same intensity as a point on the x axis with x-polarized light.

With the excitation field determined for both components of the light illuminating the lens, it is possible to determine the excitation field for an arbitrary illumination polarization. The illumination polarization is defined by the following vector.

$$\boldsymbol{\gamma}_e = \begin{bmatrix} \gamma_{e,x} \\ \gamma_{e,y} \end{bmatrix} \tag{4.3}$$

Where $\gamma_{e,x}$ is the complex coefficient of the x-polarized component and $\gamma_{e,y}$ is the complex coefficient of the y-polarized component. When the lens is illuminated with a plane wave of polarization $\boldsymbol{\gamma}_e$ the following electric field is produced.

$$\vec{h}_e(\mathbf{r}) = \gamma_{e,x}\, \vec{a}_x(\mathbf{r}; \lambda_e) + \gamma_{e,y}\, \vec{a}_y(\mathbf{r}; \lambda_e) \tag{4.4}$$

This serves as the excitation field $\vec{h}_e(\mathbf{r})$ as used in Section 2.1.

Excitation from an Opposing Direction

For opposing-lens systems such as 4Pi and I^{5}M, it is necessary to calculate the field from an opposing lens. In these systems light from a single source traverses two paths to two opposing objectives. These objectives are positioned such that they have a common focal spot and thus are illuminated from opposing directions. It will be assumed that the illumination light travels the two paths in the $y - z$ plane (and positive z is along the first lens' optic axis). This means that light that is x-polarized at the source will be in phase between the two lenses (assuming the path lengths are exactly equal) as it is perpendicular to the physical rotation. Light that is y-polarized at the source will be 180° out of phase as each has been physically rotated to get to its respective lens. This is shown in Figure 4·2.

The following simple transformation describes this process of flipping the lens illumination direction.

$$T_f = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \quad T_f^{-1} = T_f \tag{4.5}$$
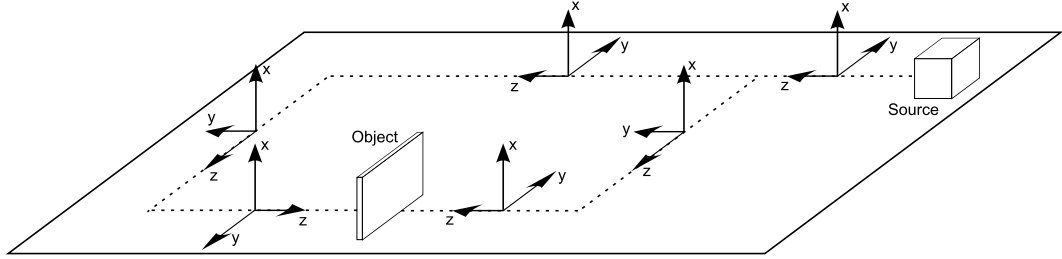
Figure 4·2: Orientations of the coordinate axes in the arms of a 4Pi microscope. Note that phase changes at the mirrors have been neglected and must be included explicitly in the model. At the two objective lenses the directions of the y and z axes are opposite.

This allows the excitation pattern from the opposing lens to be defined. It is simply a matter of applying these transformations in the same way as in Equation 4.2.

$$\vec{a}_{x,f}(\mathbf{r}; \lambda_e) = T_f\, \vec{a}_x(T_f\mathbf{r}; \lambda_e) \tag{4.6}$$

$$\vec{a}_{y,f}(\mathbf{r}; \lambda_e) = T_f\, \vec{a}_y(T_f\mathbf{r}; \lambda_e) \tag{4.7}$$

4.2.2 Detection Through a Lens

It was just shown how the excitation pattern $\vec{h}_e(\mathbf{r})$ for a single lens can be calculated from the amplitude spread function. A similar result for the detection pattern $\vec{h}_d(\mathbf{r}; \lambda_d)$ will now be shown.

The Field at the Detector Plane

While it was assumed that the objective lens was illuminated with a plane wave, the emitted light must be focused on to a detector. It will be assumed that the objective lens and a low NA tube lens form a 4f system (Saleh and Teich, 1991) between the focal plane and the detector plane. This means that each point in the object is imaged to a pattern on the detector plane. This imaging system is depicted in Figure 4·3. The detection pattern defines the imaging properties of this system — i.e. the relation between a point in the object plane and the image on the detector plane. Note that there will be two detection patterns — one that describes the generation of an x-polarized point in the detector plane
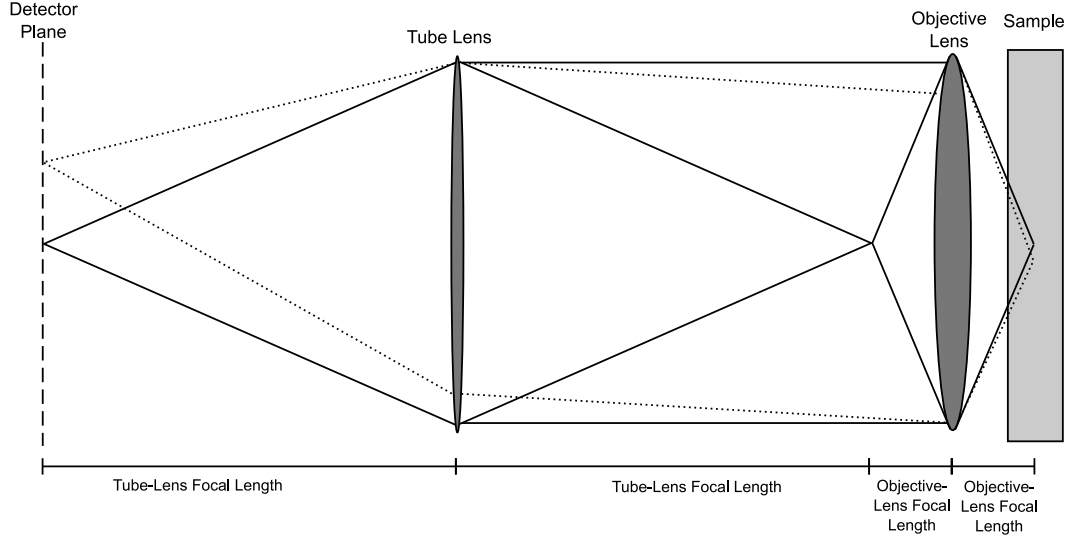
Figure 4·3: 4f microscopy detection apparatus. Each point in the sample is focused to a point in the detection plane with a magnification defined by the lenses' focal lengths.

and one that describes a y-polarized point. The detector plane location is determined by the focal lengths as shown in Figure 4·3.

It will be assumed that a source offset from the detector origin (which is the intersection of the detector plane and the optic axis) by $\mathbf{r}_d = (x_d, y_d, 0)$ (in demagnified units) gives a plane wave incident on the objective lens. This in turn gives the excitation pattern described in Section 4.2.1 but shifted by $\mathbf{r}_d$. So, an x-polarized source at position $\mathbf{r}_d$ results in a field $\vec{a}_x(\mathbf{r} - \mathbf{r}_d; \lambda)$ in the vicinity of the focus while a y-polarized source gives $\vec{a}_y(\mathbf{r} - \mathbf{r}_d; \lambda)$. This relation defines the Green's function (Barton, 1989) $g(\mathbf{r}, \mathbf{r}_d; u, v)$ where u takes on values x, y, z and represents the polarization at $\mathbf{r}$ and v represents the polarization at $\mathbf{r}_d$.

$$g(\mathbf{r}, \mathbf{r}_d; u, v) = [\vec{a}_v(\mathbf{r} - \mathbf{r}_d; \lambda)]_u \qquad (4.8)$$

The Rayleigh-Carson reciprocity theorem (Lorentz, 1896; Carson, 1924; Potton, 2004) can then be applied to calculate the field at the detector generated by a source near the focus. Rayleigh-Carson reciprocity is equivalent to the following symmetry condition on the Green's function, and applies in systems such as this where the materials used are

linear and the permittivity and permeability tensors are symmetric (Landon et al., 1984; Kong, 1990).

$$g(\mathbf{r}, \mathbf{r}_d; u, v) = g(\mathbf{r}_d, \mathbf{r}; v, u) \tag{4.9}$$

Physically, this symmetry corresponds to swapping the positions of the source and the measured field. Thus the known relation between a source at the detector plane and the field in focal region can be used to find how a source in the focal region affects the field at the detector plane. Mathematically, applying the Green's function allows the field at the detector plane $\vec{\chi}(\mathbf{r}_d; \lambda_d)$ to be calculated from the source distribution at wavelength λ_d $\vec{\xi}(\mathbf{r}; \lambda_d)$ which is itself defined by the fluorescent object.

$$
\begin{aligned}
[\vec{\chi}(\mathbf{r}_d; \lambda_d)]_x &= \sum_{u=x,y,z} \int [\vec{\xi}(\mathbf{r}; \lambda_d)]_u \, g(\mathbf{r}_d, \mathbf{r}; x, u) d\mathbf{r} \\
&= \int \vec{\xi}(\mathbf{r}; \lambda_d) \cdot \vec{a}_x(\mathbf{r} - \mathbf{r}_d; \lambda_d) d\mathbf{r} \tag{4.10}
\end{aligned}
$$

$$
\begin{aligned}
[\vec{\chi}(\mathbf{r}_d; \lambda_d)]_y &= \sum_{u=x,y,z} \int [\vec{\xi}(\mathbf{r}; \lambda_d)]_u \, g(\mathbf{r}_d, \mathbf{r}; y, u) d\mathbf{r} \\
&= \int \vec{\xi}(\mathbf{r}; \lambda_d) \cdot \vec{a}_y(\mathbf{r} - \mathbf{r}_d; \lambda_d) d\mathbf{r} \tag{4.11}
\end{aligned}
$$

The above equations are written in terms of the x- and y-polarization at the detector. It has been assumed that there is no significant z-polarization component, as the field produced by the low-NA tube lens has low curvature.

These equations represent an inner product in linear algebra. It consists of the dot product of the two vector functions integrated over space. Note that when $\vec{\xi}(\mathbf{r}, \lambda_d)$ matches the conjugate of the excitation pattern (at the detection wavelength), the projection reaches its maximum value. This corresponds to the excitation pattern traveling in the other direction — i.e. the lens detection pattern is matched to it's counter-propagating excitation pattern. This is an intuitively satisfying result. By examining the expressions for $\vec{a}_x(\mathbf{r}; \lambda_d)$ in (Richards and Wolf, 1959), it is easy to see that $\vec{a}_x(\mathbf{r}; \lambda_d)$ and $\vec{a}_y(\mathbf{r}; \lambda_d)$ are orthogonal using this inner product. This is certainly desirable as between the objective and tube lens the two components of Equation 4.10 and Equation 4.11 are modeled as x- and y-polarized

plane waves traveling in the same direction — these are also orthogonal. It would be expected that the modes would remain orthogonal throughout the system.

Measuring Intensity Over a Finite Detector Area

The principle of confocal detection relies on measuring the emitted light with a small detector. In many papers, e.g. (Nagorni and Hell, 2001b), this detector is modeled as infinitely small. The effect of a finite-sized detector has been investigated (Gu and Sheppard, 1992) and will be included in the analysis here.

It will be assumed that the detector operation can be modeled as integrating the intensity of the field over the detector plane after it has been multiplied by some detector profile. The intensity function is $|\vec{\chi}(\mathbf{r}_d; \lambda_d)|^2$ and the detector profile (which is real, non-negative and two dimensional in the detector plane) is $s(\mathbf{r}_d)$. The measured output at the detection wavelength $d(\lambda_d)$ is given by the following.

$$d(\lambda_d) = \int s(\mathbf{r}_d) \left\{ |[\vec{\chi}(\mathbf{r}_d; \lambda_d)]_x|^2 + |[\vec{\chi}(\mathbf{r}_d; \lambda_d)]_y|^2 \right\} d\mathbf{r}_d \qquad (4.12)$$

The data has been written as a continuous function of λ_d and without noise. Discretization and noise issues are assumed to be understood and treated as in Chapter 2. A simple expression for the detection profile $s(\mathbf{r}_d)$ would be one that is 1 within the detector's physical aperture and 0 outside.

4.2.3 Using Statistical Methods to Extract a Linear Model

The relations in Equation 4.10, Equation 4.11 and Equation 4.12 connect the collected data $d(\lambda_d)$ to the source distribution $\vec{\xi}(\mathbf{r}; \lambda_d)$. The model that is required is one that relates the collected data to the fluorescent object $o(\mathbf{r})$. Such a model can be found by incorporating the statistics of the emitted light and making certain assumptions.

The analysis of Section 4.2.2 implicitly assumed a deterministic, monochromatic source field was emitted from the fluorophores. In reality the emitted field is broadband and stochastic. The optical spectral detector will measure the time-averaged intensity at wave-

length λ_d as it responds slowly in comparison to optical time scales. As in Section 2.1, ergodicity will be invoked to write this time average as an expectation.

$$
\begin{aligned}
d(\lambda_d) & = E\left\{\int s(\mathbf{r}_d)\left\{|[\vec{\chi}(\mathbf{r}_d; \lambda_d)]_x|^2 + |[\vec{\chi}(\mathbf{r}_d; \lambda_d)]_y|^2\right\} d\mathbf{r}_d\right\} \\
& = \int s(\mathbf{r}_d)E\left\{|[\vec{\chi}(\mathbf{r}_d; \lambda_d)]_x|^2\right\} d\mathbf{r}_d + \int s(\mathbf{r}_d)E\left\{|[\vec{\chi}(\mathbf{r}_d; \lambda_d)]_y|^2\right\} d\mathbf{r}_d \quad (4.13)
\end{aligned}
$$

The statistical analysis employed here should not be confused with the Poisson statistics of the noise. The expectation in Equation 4.13 is used in place of a time average to calculate a deterministic (i.e. noise-free) data function. The resulting $d(\lambda_d)$ expression is indeed the mean of the data if detection noise is considered but the expectation used in the equation above is over uncertainty in the fluorophore emission, not uncertainty in the detection process.

The deterministic data is thus the integral of the expected power over the detector in an x-polarized orientation plus the integral of the expected power in a y-polarized orientation. The following analysis uses Equation 4.10 to show how the x-polarized component can be calculated. The y-polarized component follows a similar derivation.

$$
\begin{aligned}
E\left\{|[\vec{\chi}(\mathbf{r}_d; \lambda_d)]_x|^2\right\} & = E\left[\left|\int \vec{\xi}(\mathbf{r}; \lambda_d) \cdot \vec{a}_x(\mathbf{r} - \mathbf{r}_d; \lambda_d)d\mathbf{r}\right|^2\right] \\
& = E\left[\int \vec{\xi}(\mathbf{r}; \lambda_d) \cdot \vec{a}_x(\mathbf{r} - \mathbf{r}_d; \lambda_d)d\mathbf{r} \times \right. \\
& \qquad \left. \left(\int \vec{\xi}(\mathbf{r}'; \lambda_d) \cdot \vec{a}_x(\mathbf{r}' - \mathbf{r}_d; \lambda_d)d\mathbf{r}'\right)^*\right] \\
& = \int\int \sum_{u=x,y,z} \sum_{v=x,y,z} E\left\{[\vec{\xi}(\mathbf{r}; \lambda_d)]_u[\vec{\xi}(\mathbf{r}'; \lambda_d)]_v^*\right\} \times \\
& \qquad [\vec{a}_x(\mathbf{r} - \mathbf{r}_d; \lambda_d)]_u[\vec{a}_x(\mathbf{r}' - \mathbf{r}_d; \lambda_d)]_v^* \, d\mathbf{r}\, d\mathbf{r}' \quad (4.14)
\end{aligned}
$$

A spectral-domain, vectorial version of Equation 2.1 will now be used. The assumptions implicit in this equation are that the emission power from a given point is proportional to the fluorophore density, the spectral envelope amplitude at the emission wavelength and the excitation intensity at that point. Also, the emissions from different spatial positions are uncorrelated, isotropically radiating and produce unpolarized light. Note that in appli-

cations such as single-molecule detection (Basché et al., 1997) the emitted light will have preferential polarization states and emission directions depending on the orientation of the molecule. The expression below essentially assumes emission from a significant number of randomly oriented molecules. If this was not the case modifications would have to be made, e.g. in (Moiseev et al., 2004a) a SSFM system is modeled assuming a dipole emitter.

$$E\{[\vec{\xi}(\mathbf{r};\lambda_d)]_u[\vec{\xi}(\mathbf{r}';\lambda_d)]_v^*\} \propto \epsilon(\lambda_d)|\vec{h}_e(\mathbf{r})|^2 o(\mathbf{r})\delta(\mathbf{r}-\mathbf{r}')\delta_{uv} \tag{4.15}$$

Note that the second delta above is the Kronecker delta while first is the Dirac delta function.

Using the above relation, Equation 4.14 simplifies.

$$\begin{aligned} E\left\{|[\vec{\chi}(\mathbf{r}_d;\lambda_d)]_x|^2\right\} &= \int \sum_{u=x,y,z} \epsilon(\lambda_d)|\vec{h}_e(\mathbf{r})|^2 o(\mathbf{r})[\vec{a}_x(\mathbf{r}-\mathbf{r}_d;\lambda_d)]_u[\vec{a}_x(\mathbf{r}-\mathbf{r}_d;\lambda_d)]_u^* \, d\mathbf{r} \\ &= \int \epsilon(\lambda_d)|\vec{h}_e(\mathbf{r})|^2 |\vec{a}_x(\mathbf{r}-\mathbf{r}_d;\lambda_d)|^2 o(\mathbf{r}) \, d\mathbf{r} \end{aligned} \tag{4.16}$$

The same steps can be used to give the following for the y-polarized component.

$$E\left\{|[\vec{\chi}(\mathbf{r}_d;\lambda_d)]_y|^2\right\} = \int \epsilon(\lambda_d)|\vec{h}_e(\mathbf{r})|^2 |\vec{a}_y(\mathbf{r}-\mathbf{r}_d;\lambda_d)|^2 o(\mathbf{r}) \, d\mathbf{r} \tag{4.17}$$

Taking these back into the detection relation of Equation 4.13 gives the observation equation for a single lens.

$$d(\lambda_d) = \epsilon(\lambda_d) \int s(\mathbf{r}_d) \int |\vec{h}_e(\mathbf{r})|^2 \left[|\vec{a}_x(\mathbf{r}-\mathbf{r}_d;\lambda_d)|^2 + |\vec{a}_y(\mathbf{r}-\mathbf{r}_d;\lambda_d)|^2\right] o(\mathbf{r}) \, d\mathbf{r} \, d\mathbf{r}' \tag{4.18}$$

Terms can then be collected to define a detection pattern $|\vec{h}_d(\mathbf{r};\lambda_d)|^2$.

$$|\vec{h}_d(\mathbf{r};\lambda_d)|^2 = \int s(\mathbf{r}_d) \left[|\vec{a}_x(\mathbf{r}-\mathbf{r}_d;\lambda_d)|^2 + |\vec{a}_y(\mathbf{r}-\mathbf{r}_d;\lambda_d)|^2\right] \, d\mathbf{r}_d \tag{4.19}$$

This gives the following model relating the object to the data. It is analogous to the

observation equation for SSFM (Equation 2.5) but derived for fully vectorial fields.

$$d(\lambda_d) = \epsilon(\lambda_d) \int |\vec{h}_e(\mathbf{r})|^2 |\vec{h}_d(\mathbf{r}; \lambda_d)|^2 o(\mathbf{r}) \, d\mathbf{r} \tag{4.20}$$

4.3 Point Spread Functions and Optical Transfer Functions

A linear model relating the object to the data point collected at λ_d is given by Equation 4.20. However, most microscopy systems collect multiple data points at any wavelength measured. This imaging usually involves mapping each position in the object to a given data point. This can either be done by scanning through the object and collecting a data point at each offset (as in confocal or 4Pi microscopy) or by simultaneous collection over the whole detector plane (as in widefield microscopy). A combination of these methods may also be used as in three dimensional widefield microscopy where only the axial dimension is scanned. All of these cases involve a simple modification to the model of Equation 4.20.

$$
\begin{aligned}
d(\mathbf{r}_s; \lambda_d) &= \epsilon(\lambda_d) \int |\vec{h}_e(\mathbf{r})|^2 |\vec{h}_d(\mathbf{r}; \lambda_d)|^2 o(\mathbf{r} - \mathbf{r}_s) \, d\mathbf{r} \\
&= h(\mathbf{r}_s; \lambda_d) *_{\mathbf{r}_s} o(\mathbf{r}_s)
\end{aligned} \tag{4.21}
$$

The data is generally a function of the spatial variable $\mathbf{r}_s$ and $\mathbf{r}_s$ can be regarded as a scanning shift applied to the object. In reality this scanning offset may actually come from the apparatus (in which case the shift would appear in the excitation and detection patterns) or from some form of widefield detection. All these cases are mathematically equivalent to Equation 4.21 which can be recognized as a convolution integral. The point spread function (at λ_d) of this integral is given below.

$$h(\mathbf{r}; \lambda_d) = \epsilon(\lambda_d) |\vec{h}_e(-\mathbf{r})|^2 |\vec{h}_d(-\mathbf{r}; \lambda_d)|^2 \tag{4.22}$$

Note that the scanning offset $\mathbf{r}_s$ will generally be a discrete set of points but it will be treated as a continuous variable for ease of analysis. Again, the sampling issues discussed in Chapter 2 can be used to relate the continuous and discrete data sets. Sampling issues

are also addressed further in Chapter 6.

Now that the microscopy system is modeled in the language of standard linear, time-invariant systems, the extensive work on this class of systems can be leveraged. In particular the OTF $H(\mathbf{k})$ can be calculated by taking the Fourier transform of the point spread function. As discussed in detail in the next chapter, the support of the OTF determines the maximum achievable resolution of the system. The OTF relates the data to the object in the Fourier domain as shown below.

$$D(\mathbf{k}; \lambda_d) = H(\mathbf{k}; \lambda_d)O(\mathbf{k}) \tag{4.23}$$

Specific microscopy systems will now be examined using this framework.

4.3.1 Widefield Microscopy

In widefield microscopy, the sample is illuminated with a uniform field and detected through a lens as in Equation 4.19. The PSF for a widefield system operating at a detection wavelength of 530nm was calculated. The objective lens used has a numerical aperture of 1.35 and is used in a medium with a refractive index of 1.51. Circularly polarized excitation light and a point detector are considered. These parameters define a system operating at visible wavelengths and near the practical limits of achievable resolution capability. The relevant electric fields were calculated on a $4\mu m \times 4\mu m \times 8\mu m$ (x, y, z) volume with $128 \times 128 \times 256$ sample points.

Once the excitation and detection patterns have been calculated, they are simply multiplied, as in Equation 4.22, to get the point spread function. These functions and the relations between them can also be examined in the Fourier domain — a Fast Fourier Transform (FFT) algorithm (Cooley and Tukey, 1965; Duhamel and Vetterli, 1990) is used to perform a Fourier Transform numerically. The widefield PSFs are shown in Figure 4·4 as are the OTF functions. Since excitation and detection PSFs are multiplied to give the total PSF, the excitation and detection OTFs are convolved to give the total OTF. In this case the excitation OTF is a delta function so the total OTF is the same as the detection OTF.
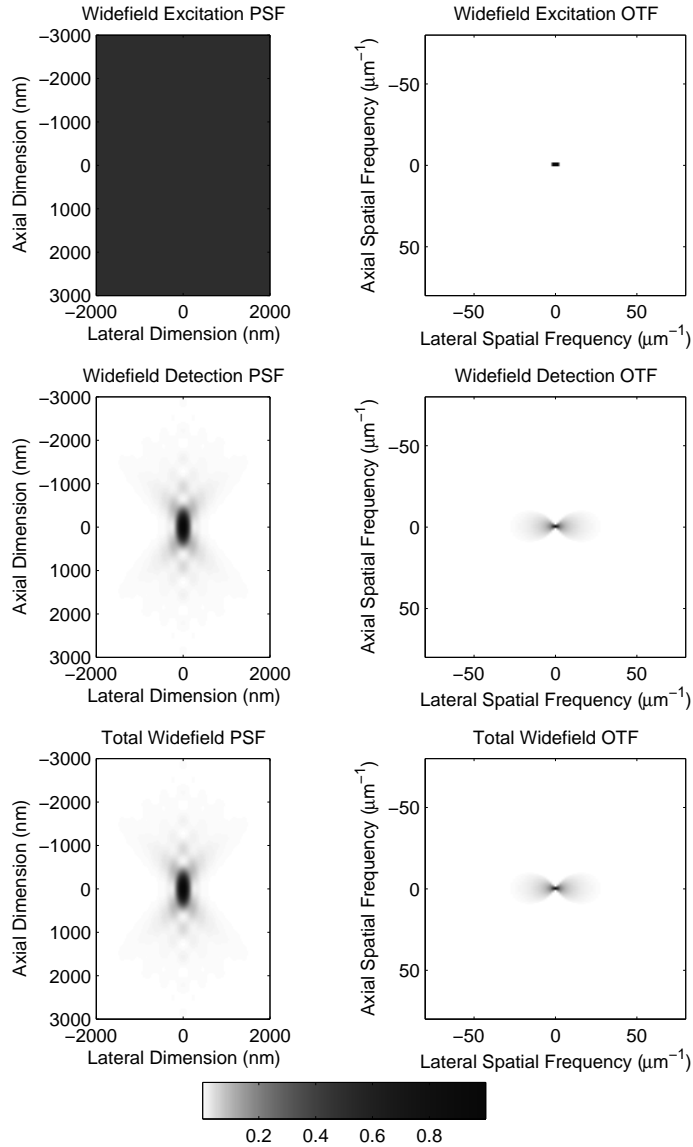
Figure 4·4: Point spread functions and optical transfer functions for a widefield microscope. The PSF for the excitation light and the PSF for the detection light are shown while the total PSF is the product of these two as in Equation 4.22. All figures are plotted on a non-linear scale to show low-level detail and all functions are normalized to have a maximum value of 1. Note that these plots are two-dimensional slices of three-dimensional functions that are all cylindrically symmetric about the z axis. The excitation OTF has a single non-zero pixel — this has been artificially broadened so it is visible in the plot.

The OTF is a very useful analytic tool as it shows how strongly each spatial frequency in the object appears in the data.

4.3.2 Confocal Microscopy

Confocal microscopy involves focusing both the excitation and detection light. Mathematically, the excitation pattern $\vec{h}_e(\mathbf{r})$ is defined by Equation 4.4 and the detection pattern $\vec{h}_d(\mathbf{r}; \lambda_d)$ is defined by Equation 4.19. While the focused excitation does necessitate three-dimensional scanning in order to build a full image, it has a significant advantage in terms of resolution. A confocal PSF is calculated here and has also been examined in the literature (van der Voort and Brakenhoff, 1990) using a similar methodology. The excitation wavelength used is 488nm while all other system parameters are the same as given in the widefield example of Section 4.3.1. The results are shown in Figure 4·5.

It can be seen that taking the product of two focused fields means that confocal microscopy has a smaller PSF than the widefield instrument. This means a greater resolution as indicated by the OTF which now has a greater extent.

4.3.3 4Pi Type C Microscopy

The point spread function for a 4Pi microscope can be calculated in the same manner as the confocal PSF was above. For a Type C system, the single-lens amplitude spread function is simply replaced by the pattern resulting from two opposing lenses with a common focal point (which is the origin of the coordinate system). The 4Pi excitation pattern is given below (c.f. Equation 4.4).

$$\vec{h}_{e,4\mathrm{Pi}}(\mathbf{r}) = \gamma_{e,x}\left[\vec{a}_x(\mathbf{r}; \lambda_e) + \vec{a}_{x,f}(\mathbf{r}; \lambda_e)\right] + \gamma_{e,y}\left[\vec{a}_y(\mathbf{r}; \lambda_e) + \vec{a}_{y,f}(\mathbf{r}; \lambda_e)\right] \tag{4.24}$$

Similarly, the following equation describes the detection pattern.

$$|\vec{h}_{d,4\mathrm{Pi}}(\mathbf{r}; \lambda_d)|^2 = \int s(\mathbf{r}_d)\left[|\vec{a}_x(\mathbf{r}; \lambda_d) + \vec{a}_{x,f}(\mathbf{r}; \lambda_d)|^2 + |\vec{a}_y(\mathbf{r}; \lambda_d) + \vec{a}_{y,f}(\mathbf{r}; \lambda_d)|^2\right] d\mathbf{r}_d \tag{4.25}$$

Calculating the total PSF is simply a matter of applying Equation 4.22.
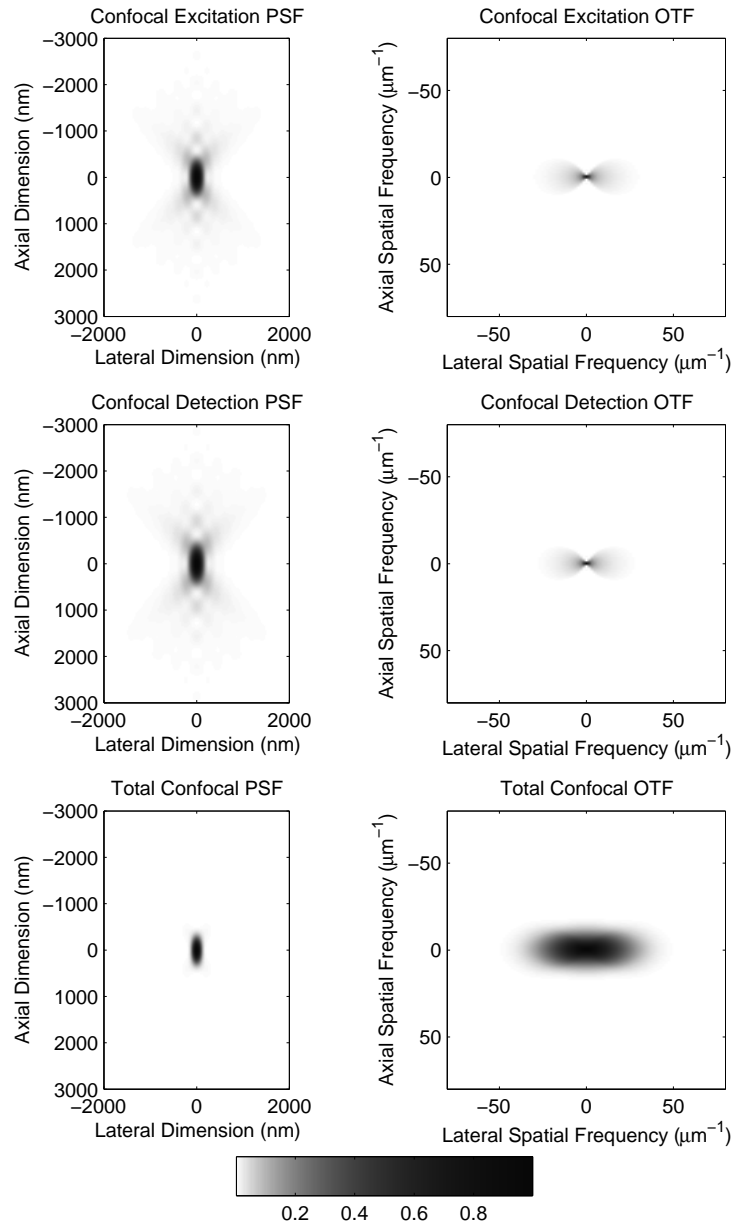
Figure 4·5: Point spread functions and optical transfer functions for a confocal microscope. The PSF for the excitation light and the PSF for the detection light are shown while the total PSF is the product of these two as in Equation 4.22. All figures are plotted on a non-linear scale to show low-level detail and all functions are normalized to have a maximum value of 1. Note that these plots are two-dimensional slices of three-dimensional functions that are all cylindrically symmetric about the z axis.

Example PSFs were again calculated. As in the confocal example, the detection wavelength is 530nm, the excitation wavelength is 488nm, the objective lenses have a numerical aperture of 1.35 and are used in a medium with a refractive index of 1.51. Circularly polarized excitation light and a point detector are considered. The relevant electric fields were calculated on a $4\mu m \times 4\mu m \times 8\mu m$ (x, y, z) volume with $128 \times 128 \times 256$ sample points. The resulting PSFs and OTFs are shown in Figure 4·6. It should also be noted that the reflections used to produce a counter-propagating beam are assumed to give zero phase difference between the y-polarized fields. The situation illustrated in Figure 4·2 introduces $180°$ phase shift. A zero phase shift is used as it means that the x- and y-polarized components exhibit the same interference effects. The possibility of the 4Pi setup producing different phases for the x- and y-polarized components has been recognized and explored in the literature (Bahlmann and Hell, 2000).

The effects of interference can be clearly seen in Figure 4·6 — a standing-wave structure is now present in the PSFs and the OTFs exhibit axial side bands. These result in an OTF that is significantly extended in the axial direction. This allows a higher axial resolution to be achieved but only after post-processing. The sidelobes on the PSF will result in ringing effects on the raw data. In a well designed system the appropriate image processing can remove this ringing effect to give a clear, high-resolution image (Schrader et al., 1998b; Nagorni and Hell, 1998; Nagorni and Hell, 2001a).

Typically only a single wavelength is detected in 4Pi systems, unlike SSFM. Measurements of the 4Pi PSF and OTF have been taken by measuring point-like objects (Schrader et al., 1997) and shown to agree well with the theory presented here. This model also assumes that there are no significant object induced distortions — that is, the PSF is assumed unaffected by interactions with the object. This assumption is also implicit for the confocal calculation but is more significant in 4Pi microscopy where relatively minor phase distortions can significantly alter the PSF due to altered interference effects. This is a major issue in 4Pi microscopy as many samples of interest will induce phase distortions. While phase changes can be estimated (Hell et al., 2002; Blanca et al., 2002) and corrected
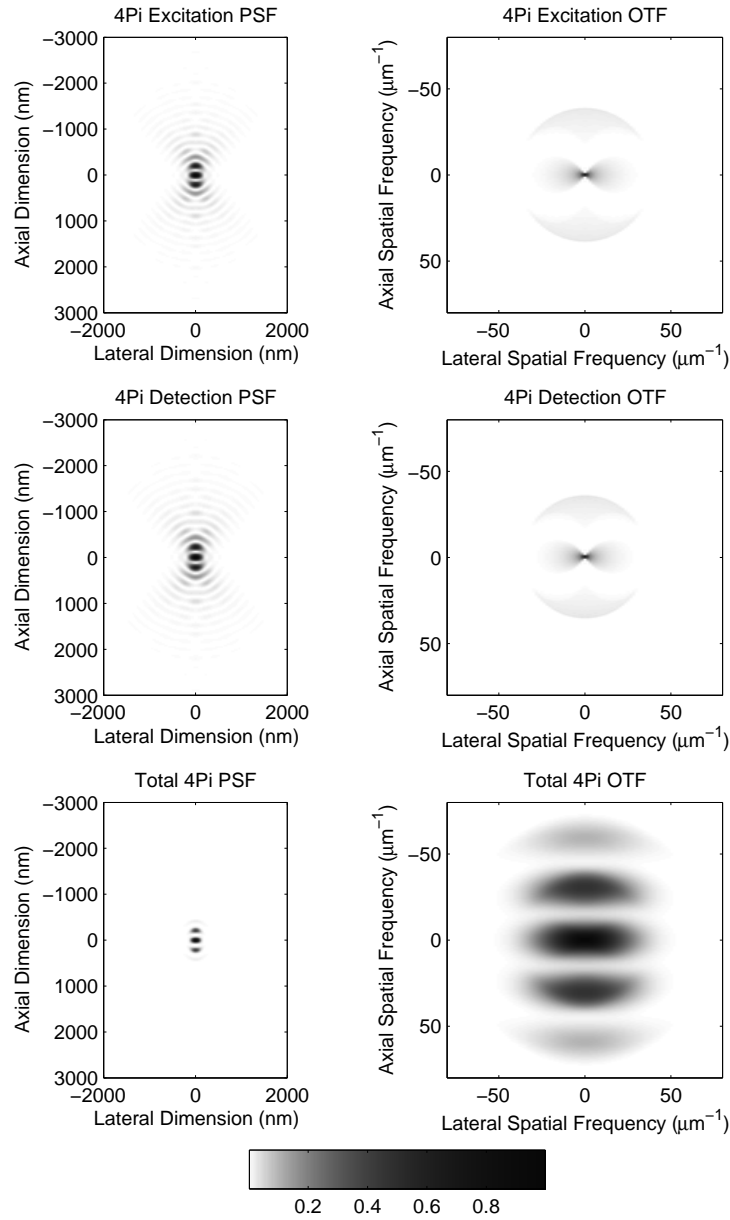
Figure 4·6: Point spread functions and optical transfer functions for a 4Pi Type C microscope. The PSF for the excitation light and the PSF for the detection light are shown while the total PSF is the product of these two as in Equation 4.22. All figures are plotted on a non-linear scale to show low-level detail and all functions are normalized to have a maximum value of 1. Note that these plots are two-dimensional slices of three-dimensional functions that are all cylindrically symmetric about the z axis.

to some extent (Egner et al., 1998; Schrader et al., 1998a), it is important to note that phase changes may cause the PSF to vary across the object (giving a system that is not, strictly speaking, linear and shift-invariant). This limits the utility of a simple PSF model for deconvolution purposes.

4.3.4   SSFM

As discussed in Section 4.1, the most general 4Pi/SSFM system includes spectral detection, possibly unequal lengths between the two detection/illumination paths and possibly mismatched foci between the lenses. Detection at an arbitrary wavelength has already been included in the models developed to this point. The unequal path lengths serve to add a phase delay between the paths, i.e. one of the arm's detection/excitation pattern is multiplied by a phase term $e^{i\phi}$. The focal point offset simply adds a spatial shift of $\mathbf{r}_o$ between the lens' patterns. These terms can be added to Equation 4.24 to give the following excitation model.

$$\vec{h}_{e,\mathrm{SSFM}}(\mathbf{r}) = \gamma_{e,x}\left[\vec{a}_x(\mathbf{r};\lambda_e) + e^{i\phi_e}\vec{a}_{x,f}(\mathbf{r}-\mathbf{r}_o;\lambda_e)\right] + \gamma_{e,y}\left[\vec{a}_y(\mathbf{r};\lambda_e) + e^{i\phi_e}\vec{a}_{y,f}(\mathbf{r}-\mathbf{r}_o;\lambda_e)\right]$$

$$(4.26)$$

Similarly, Equation 4.25 can be modified to give the detection model. Note that the same physical path difference will give a differing phase delay for different wavelengths.

$$\begin{aligned}
|\vec{h}_{d,\mathrm{SSFM}}(\mathbf{r},\lambda_d)|^2 &= \int s(\mathbf{r}_d)\left[\left|\vec{a}_x(\mathbf{r};\lambda_d) + e^{i\phi_d(\lambda_d)}\vec{a}_{x,f}(\mathbf{r}-\mathbf{r}_o;\lambda_d)\right|^2 + \right.\\
&\quad \left. \left|\vec{a}_y(\mathbf{r};\lambda_d) + e^{i\phi_d(\lambda_d)}\vec{a}_{y,f}(\mathbf{r}-\mathbf{r}_o;\lambda_d)\right|^2\right]\,d\mathbf{r}_d \quad (4.27)
\end{aligned}$$

Again, Equation 4.22 is applied to get the total PSF. Once the PSFs and OTFs have been calculated they can be employed in analyzing the resolution of a SSFM system. The critical question is, does collecting at multiple wavelengths, including phase delays and/or introducing foci offset improve the resolution?

Some example PSFs are shown in Figure 4·7. The simulation parameters are the same as those given in Section 4.3.3 and two detection wavelengths are used. In general a
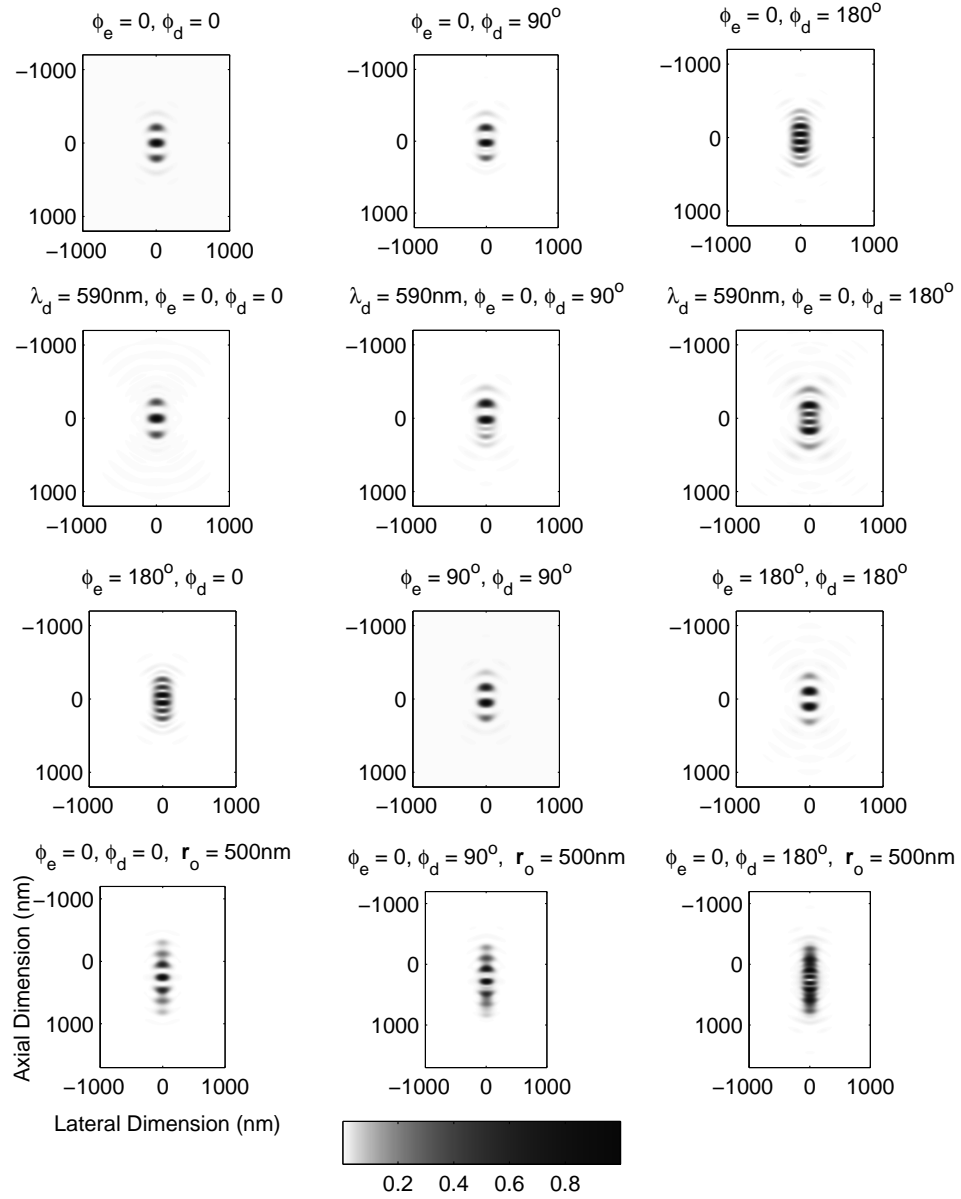
Figure 4·7: Example point spread functions for a SSFM system. All figures are plotted on a non-linear scale to show low-level detail and all functions are normalized to have a maximum value of 1. Values of the excitation and detection phases ϕ_e and ϕ_d are displayed on the graphs. The excitation wavelength λ_e is 488nm. Unless otherwise noted the detection wavelength λ_d is 510nm and the focal point offset $\mathbf{r}_o$ is $\mathbf{0}$. Note that these plots are two-dimensional slices of three-dimensional functions that are all cylindrically symmetric about the z axis.

spectrum of wavelengths will be collected but the PSFs and OTFs considered here are for two illustrative wavelengths from either end of the range collected in the examples of Chapter 3. Note that only the total PSFs (and not the excitation and detection cases) are shown for brevity. Three of the displayed cases include an offset of 500nm between the two lenses. The system used in (Swan et al., 2003) places the object several microns above the mirror and thus corresponds to an offset much higher than 500nm. However, that particular system used low NA lenses — for the high NA case considered here, such a large offset would result in very little interaction between the two lenses due to the short depth of focus and for that reason a lesser offset is considered.

It can be seen that the phase introduced by the path-length difference affects the PSFs by changing the lobe positions and shapes. There are also noticeable differences at different detection wavelengths. As would be expected, an offset between the two lenses results in an extended PSF. These plots show that SSFM does indeed significantly alter the 4Pi PSF but it is not clear how this affects resolution.

The OTFs corresponding to the PSFs of Figure 4·7 are shown in Figure 4·8. Note that in these plots the OTF magnitude is displayed — unlike the previously considered instruments, asymmetries in the PSFs result in complex OTFs. The OTF plots have a strong connection to the resolution of the system. Better OTFs are those that have a higher magnitude over a greater range of spatial frequencies. The 'best' OTF by these criteria is arguably the case where $\lambda_d =$510nm and $\phi_e = \phi_d = 0$. When a longer detection wavelength is used a slightly more compact OTF is observed. This makes sense as a shorter wavelength gives a smaller PSF and hence a larger OTF support. The minimum viable detection wavelength is determined by the fluorophore's emission properties. Non-zero phase terms have a varying effect on the OTFs but do not improve the OTF extent. Destructive interference, rather than the standard constructive case, in 4Pi microscopy has been addressed in the literature (Hell and Nagorni, 1998) and shown to not produce a large change in image quality. It can also be seen that an offset between the foci does not improve the OTF. This is also intuitive since the offset will tend to broaden the PSF.
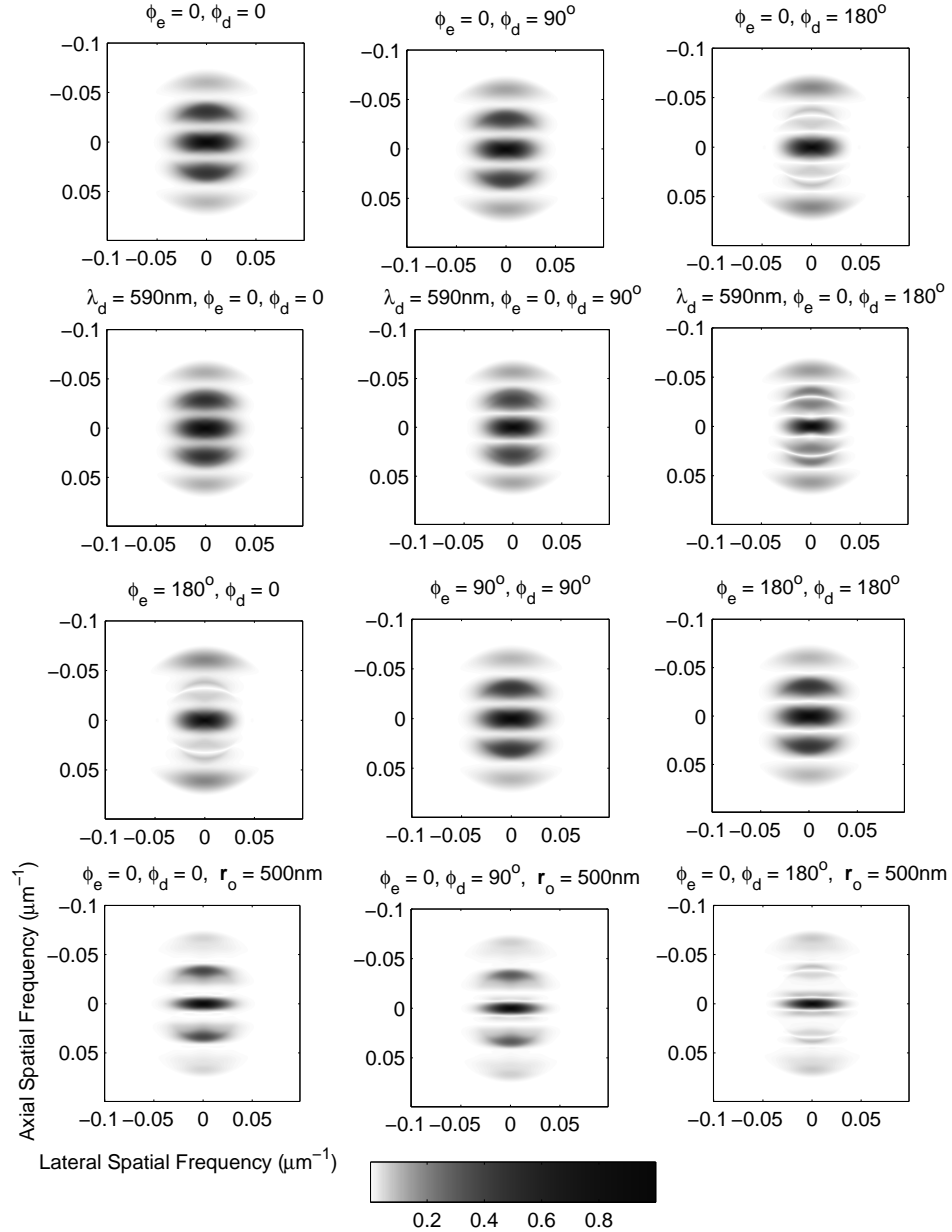
Figure 4·8: Example optical transfer function magnitudes for a SSFM system. All figures are plotted on a non-linear scale to show low-level detail and all functions are normalized to have a maximum value of 1. Values of the excitation and detection phases ϕ_e and ϕ_d are displayed on the graphs. The excitation wavelength λ_e is 488nm. Unless otherwise noted the detection wavelength λ_d is 510nm and the focal point offset $\mathbf{r}_o$ is $\mathbf{0}$. Note that these plots are two-dimensional slices of three-dimensional functions that are all cylindrically symmetric about the z axis.

It also means that a given point in the sample cannot be simultaneously in focus in both lenses. For these reasons, offset foci will not be considered in the rest of this work.

The example OTFs calculated for the SSFM system indicate that the additional SSFM features (spectral detection, unequal path lengths, foci offset) do not improve the system's resolution relative to a 4Pi system with the same operating parameters. This is based on the fact that the SSFM system can be regarded as a series of linear shift-invariant systems operating in parallel. Each wavelength detected represents a single system. The addition of phase delays and/or foci offset does not improve the OTF at a given wavelength. This means that spatial frequencies not passed by the 4Pi system are not passed by the SSFM system either. This is why the resolution is not improved. The multi-channel observations may give advantages other than improved resolution over a single-channel system and this idea is explored further in later chapters.

4.4   Summary

Comparing the apparatus used to build a 4Pi system to that used in a SSFM system reveals significant similarities. A scanning SSFM system can in fact be regarded as a generalization of a 4Pi system despite differences in the practical aspects of microscope construction and operation. Recognizing commonalities between the two systems is very useful for theoretical analysis as it means the SSFM system can be analyzed as a linear shift-invariant system (shift-invariance was not used in Chapter 3 as scanning was not included). As was done here, a linear systems analysis is commonly used to compare the performance of different microscopy systems (Gustafsson, 1999; Nagorni and Hell, 2001b).

A method for calculating point spread functions was then described. This model is fully polarimetric and does not employ commonly-used low-angle approximations that are incompatible with high NA lenses (Sheppard and Matthews, 1987). Using this framework, PSFs were calculated for widefield, confocal, 4Pi and SSFM systems and example plots shown. More importantly for resolution analysis, the PSFs can be used to calculate optical transfer functions which give a clear idea of which spatial frequencies are passed by the

system. The OTFs produced by the SSFM instrument were seen to be comparable or worse than those given by a 4Pi system.

There are several key assumptions involved in forming the PSF/OTF model. The first is that the imaging is being performed under optimal conditions — no abberations are present on the lens(es), the illumination power doesn't fluctuate, the fluorophore response is linear, etc. A discussion of such practical issues in confocal microscopy is given in (Pawley, 2000). The effects of noise have also been neglected in this chapter (later chapters will discuss noise in detail). Finally, the model developed has assumed that the object imaged does not affect the electric fields used to build the PSF. This approximation is only valid for thin/low-absorbing samples with relatively uniform index of refraction. While these approximations are significant, they are standard in the literature cited in this chapter and generally these methods will only be applied to samples where the assumptions apply.

This chapter indicates that when a more comprehensive model is used for SSFM, the resolution is still comparable to a 4Pi system. It is also important to realize that, up to this point, only examples have been used to characterize the SSFM system. In the next chapter a comprehensive Fourier-domain analysis is used to characterize SSFM systems in the general case. Later chapters address how the set of multiple channel measurements can be processed to improve the imaging performance, if not the resolution.

Chapter 5

Analysis of Optical Transfer Function Support

This chapter contains a general result for far-field light microscopy that has, in part, been published as a single journal paper (Davis et al., 2004c). It will be shown that the OTFs of far-field focusing microscopes are zero outside a certain support and a method for calculating this support is given. The result is based on optical transfer function analysis similar to that presented in the previous chapter. Since it is a general result, those readers interested only in the development of SSFM can probably skip this chapter. The chapter does however contain a more rigorous reasoning as to why the additional features of SSFM do not increase the support of the 4Pi OTF.

In the previous chapter it was seen that the PSFs of far-field, focusing microscopy systems can be constructed from various amplitude spread functions $\vec{a}_x(\mathbf{r}; \lambda)$. These methods can also be expressed in the Fourier domain as a construction of the OTF, e.g. (Gu and Sheppard, 1992; Sheppard and Gu, 1994; Gu and Sheppard, 1994; Arnison and Sheppard, 2002; Schönle and Hell, 2002). Just as the PSFs were calculated from the amplitude spread function, the OTFs can be expressed in terms of the Amplitude Transfer Function (ATF) $\vec{A}_x(\mathbf{k}; \lambda)$, which is the Fourier transform of $\vec{a}_x(\mathbf{r}; \lambda)$. The properties of the Fourier transform define the frequency domain construction from the spatial domain construction. The resulting OTFs can be seen to have a finite support (Sheppard, 1986a; Sheppard, 1986b; Gustafsson, 1999) and the OTF support determines the maximum achievable resolution of the system.

Key properties of the Fourier transform are given in the next section before the form of $\vec{A}_x(\mathbf{k}; \lambda)$ is discussed. The support of the OTF for several common microscopy systems is then calculated using these results. Several modified systems including SSFM and lenses

with pupil filters are then discussed, followed by an examination of performance in the presence of noise.

5.1 Fourier Transform Properties

This section presents some basic and well-known Fourier transform properties (Bracewell, 1999) as applied to vector functions. These properties will be used in the Fourier domain construction of microscope OTFs. Note that the properties are defined using the amplitude spread functions as examples but the results hold for arbitrary functions (with defined Fourier transforms).

Addition

Consider two three-dimensional vector functions. The Fourier transform of their sum is simply the sum of their Fourier transforms.

$$\vec{a}_x^{(1)}(\mathbf{r}; \lambda) + \vec{a}_x^{(2)}(\mathbf{r}; \lambda) \leftrightarrows \vec{A}_x^{(1)}(\mathbf{k}; \lambda) + \vec{A}_x^{(2)}(\mathbf{k}; \lambda) \tag{5.1}$$

This is a generalization of the same result for scalar functions.

Multiplication by a Constant

Multiplying a spatial-domain function by a complex constant affects the Fourier transform as multiplication by the same complex constant. Note that this constant can be a phase term of the form $e^{i\phi}$

$$K \, \vec{a}_x(\mathbf{r}; \lambda) \leftrightarrows K \, \vec{A}_x(\mathbf{k}; \lambda) \tag{5.2}$$

The multiplication-by-constant result also holds for scalar functions.

Shift in Space

A spatial shift of a function produces a modulation by a complex exponential in Fourier space.

$$\vec{a}_x(\mathbf{r} - \mathbf{r}_o; \lambda) \leftrightarrows e^{-i\mathbf{k}\cdot\mathbf{r}_o} \vec{A}_x(\mathbf{k}; \lambda) \tag{5.3}$$

Again, this result holds for scalar functions.

Orientation Flip

The transformation matrix of Equation 4.5 is used in Equation 4.6 and Equation 4.7 to define the amplitude spread function of a lens oriented in the opposite direction. It is straightforward to show that this results in the same transformation in the Fourier domain.

$$\vec{a}_{x,f}(\mathbf{r}; \lambda_e) = T_f\, \vec{a}_x(T_f\mathbf{r}; \lambda_e) \leftrightarrows T_f\vec{A}_x(T_f\mathbf{k}; \lambda) \tag{5.4}$$

For a scalar function, flipping any combination of the spatial axes causes the same flips in the Fourier-domain axes.

Polarization Rotation

The amplitude spread function for x-polarized light can be transformed into the amplitude spread function for the y-polarized light using Equation 4.1 and Equation 4.2. This transformation is simply a rotation of the coordinate systems followed by a reordering and/or negation of the function's vector components. This results in the same transformations being applied in Fourier space.

$$\vec{a}_y(\mathbf{r}; \lambda) = T_y\, \vec{a}_x(T_y^{-1}\mathbf{r}; \lambda_e) \leftrightarrows T_y\vec{A}_x(T_y^{-1}\mathbf{k}; \lambda) \tag{5.5}$$

Convolution

Convolution is well-defined for scalar functions and for the purposes of this work, the convolution of vector functions will be defined as the scalar convolution of each of the vector components. In either the scalar or vector case, the following equation defines the Fourier domain effects of convolving two functions. Note that in the vector case, the multiplication on the right hand side is component by component.

$$\vec{a}_x^{(1)}(\mathbf{r}; \lambda) *_{\mathbf{r}} \vec{a}_x^{(2)}(\mathbf{r}; \lambda) \leftrightarrows \vec{A}_x^{(1)}(\mathbf{k}; \lambda)\, \vec{A}_x^{(2)}(\mathbf{k}; \lambda) \tag{5.6}$$

It is also possible to define the convolution of a scalar function and a vector function — the scalar function is convolved with each component of the vector function. Again, the result above holds where the multiplication also consists of multiplying the scalar Fourier function with each component of the vector Fourier function.

Multiplication of Functions

Spatial-domain multiplication leads to Fourier-domain convolution as shown below. The same rules for vector/scalar function multiplication and convolution are used here.

$$\vec{a}_x^{(1)}(\mathbf{r}; \lambda) \, \vec{a}_x^{(2)}(\mathbf{r}; \lambda) \leftrightarrows \vec{A}_x^{(1)}(\mathbf{k}; \lambda) *_{\mathbf{k}} \vec{A}_x^{(2)}(\mathbf{k}; \lambda) \tag{5.7}$$

Complex Conjugate

The complex conjugate operation is often used when dealing with electromagnetic fields. The following equation describes the Fourier-domain effect of this operation and also applies to both scalar and vector functions.

$$\vec{a}_x^*(\mathbf{r}; \lambda) \leftrightarrows \vec{A}_x^*(-\mathbf{k}; \lambda) \tag{5.8}$$

Field Intensity

Another common operation on electromagnetic fields is to calculate the intensity. This operation can actually be written in terms of operations already discussed — specifically summation, multiplication of functions and conjugation.

$$|\vec{a}_x(\mathbf{r}; \lambda)|^2 = \sum_{u=x,y,z} [\vec{a}_x(\mathbf{r}; \lambda)]_u \, [\vec{a}_x^*(\mathbf{r}; \lambda)]_u \tag{5.9}$$

This leads to the spatial-Fourier relation shown below.

$$|\vec{a}_x(\mathbf{r}; \lambda)|^2 \leftrightarrows \sum_{u=x,y,z} \left[\vec{A}_x(\mathbf{k}; \lambda)\right]_u *_{\mathbf{k}} \left[\vec{A}_x^*(-\mathbf{k}; \lambda)\right]_u \tag{5.10}$$

Inner Product

The inner product is another potentially useful operation and can be defined in terms of multiplication of functions, summation and integration.

$$\int \vec{a}_x^{(1)}(\mathbf{r};\lambda) \cdot \left[\vec{a}_x^{(2)}(\mathbf{r};\lambda)\right]^* d\mathbf{r} = \int \sum_{u=x,y,z} \left[\vec{a}_x^{(1)}(\mathbf{r};\lambda)\right]_u \left[\vec{a}_x^{(2)}(\mathbf{r};\lambda)\right]_u^* d\mathbf{r} \qquad (5.11)$$

The following spatial-Fourier relation can be found from the definition above.

$$\int \vec{a}_x^{(1)}(\mathbf{r};\lambda) \cdot \left[\vec{a}_x^{(2)}(\mathbf{r};\lambda)\right]^* d\mathbf{r} = \int \vec{A}_x^{(1)}(\mathbf{k};\lambda) \cdot \left[\vec{A}_x^{(2)}(\mathbf{k};\lambda)\right]^* d\mathbf{k} \qquad (5.12)$$

The equations above are simply Parseval's theorem (Bracewell, 1999) which states that the inner product is the same whether it is calculated in the spatial domain or the Fourier domain.

5.2 The Amplitude Transfer Function

The effort expended to calculate the OTF directly in the Fourier domain (rather than using the spatial-domain constructions of the last chapter and subsequently taking the Fourier transform) is justified due to the special form of amplitude transfer function $\vec{A}_x(\mathbf{k};\lambda)$. A scalar form for this function was first calculated (McCutchen, 1964) and it was later realized that the result could be extended to vector optics (McCutchen, 2002). Similar vector formulations have been used for specific applications (Arnison and Sheppard, 2002; Schönle and Hell, 2002).

The ATF is special in that its support is limited to a spherical shell. This is expressed mathematically below.

$$\{\mathbf{k} : \vec{A}_x(\mathbf{k};\lambda) \neq \mathbf{0}\} = \left\{\mathbf{k} : |\mathbf{k}| = \frac{2\pi n}{\lambda}, \theta \leq \Theta\right\} \qquad (5.13)$$

Amplitude Transfer Function for an Aplanatic Lens
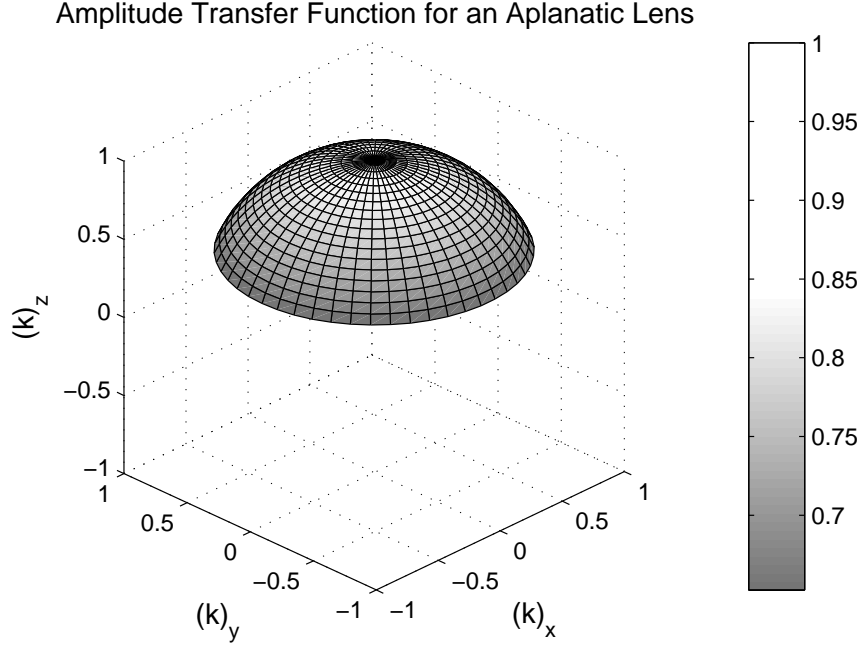
Figure 5·1: Graphical representation of the amplitude transfer function $\vec{A}(\mathbf{k};\lambda)$ for an aplanatic lens with maximum collection angle of $\Theta = 65°$. The axes have been normalized.

The angle θ represents the angle between a propagation direction and the optic (z) axis.

$$\theta = \tan^{-1}\left(\frac{\sqrt{(\mathbf{k})_x^2 + (\mathbf{k})_y^2}}{(\mathbf{k})_z}\right) \tag{5.14}$$

The limit of the angle, Θ, is the maximum collection angle of the lens. An illustration of the support of $\vec{A}_x(\mathbf{k};\lambda)$ is shown in Figure 5·1.

The support shown has an intuitive physical meaning. Each point in Fourier space represents a plane wave. The points on a sphere have a fixed $|\mathbf{k}|$ which indicates a common wavelength. The limited angle condition corresponds to the limited numerical aperture of the lens — the lens only focuses light up to a given angle. For an aplanatic lens, which is generally used for imaging, the amplitude profile across the support is given by $\sqrt{\cos\theta}$.

As will be seen in subsequent examples, the limited support of the amplitude spread function results in a limited support in the OTF for many microscopy systems. The support

of the OTF plays a critical role in the achievable resolution — the larger the support, the higher the resolution tends to be. Any spatial frequencies in the object that are outside of the support of the OTF will not have an influence on the data. This in turn implies that object components that oscillate at a frequency outside the OTF support are lost (i.e. unobserved) in the imaging. Some examples of OTF construction and calculation of the OTF support follow.

5.3 Optical Transfer Function Supports

This section contains examples of how the spatial domain construction of a PSF (as described in the last chapter) leads to a Fourier-domain construction of the OTF. This construction defines the support of the resulting OTF (Sheppard, 1986a; Sheppard, 1986b; Gustafsson, 1999).

5.3.1 Widefield Microscopy

Widefield incoherent microscopy involves a uniform illumination pattern and the detection pattern is given by Equation 4.19. Using these definitions and the Fourier domain properties described earlier in this chapter gives the following Fourier domain construction of the OTF (Sheppard and Gu, 1994; Arnison and Sheppard, 2002; Schönle and Hell, 2002).

$$
\begin{aligned}
H_{\text{Wide.}}(\mathbf{k}, \lambda_d) \;=\; S(-\mathbf{k}) \Bigg\{ &\left[\sum_{u=x,y,z} \left(\vec{A}_x(\mathbf{k};\lambda) \right)_u *_{\mathbf{k}} \left(\vec{A}_x^*(-\mathbf{k};\lambda) \right)_u \right] \\
+ &\left[\sum_{u=x,y,z} \left(\vec{A}_y(\mathbf{k};\lambda) \right)_u *_{\mathbf{k}} \left(\vec{A}_y^*(-\mathbf{k};\lambda) \right)_u \right] \Bigg\}
\end{aligned}
\tag{5.15}
$$

The equation above looks somewhat complicated but is actually just a sum of convolutions of ATF components and a flipped conjugate version of the same component. This means that each term in these sums has the same support which is therefore the support of the total OTF. Graphically, the support corresponds to taking two offset copies of the support in Figure 5·1 and for any offset where they overlap, the OTF of the widefield microscope in non-zero. The OTF for a widefield system, using a numerical aperture of 1.35, an index

Widefield Frequency Domain Support, n=1.51, NA=1.35
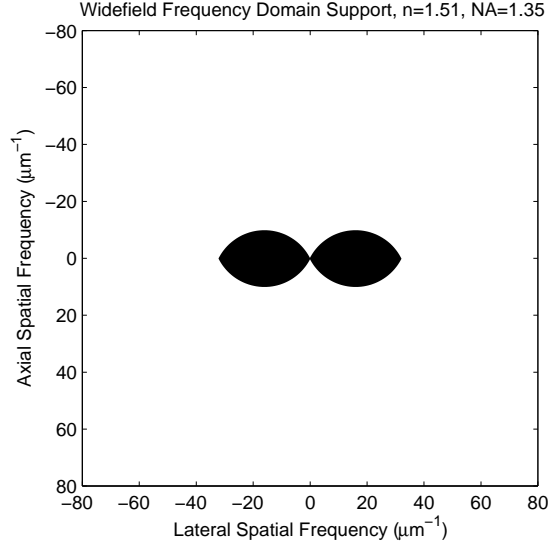
Figure 5·2: The support of the optical transfer function for a widefield microscope operating at a detection wavelength of 530nm. Note that this plot is a two-dimensional slice of the three-dimensional support that is cylindrically symmetric about the z axis.

of 1.51 and a wavelength of 530nm is shown in Figure 4·4. The support for this system is a torus shaped with a cross-section as shown in Figure 5·2.

The maximum lateral spatial frequency within the support is given by the expression below. This is easily calculated using the Fourier-domain OTF construction.

$$(\mathbf{k}_{c,\text{Wide.}})_x = (\mathbf{k}_{c,\text{Wide.}})_y = \frac{4\pi}{\lambda_d} n \sin \Theta \tag{5.16}$$

This maximum spatial frequency corresponds to spacial period of $\frac{\lambda_d}{2n \sin \Theta}$ which is the standard equation given for widefield resolution (Abbe, 1873). The maximum axial frequency is given below.

$$(\mathbf{k}_{c,\text{Wide.}})_z = \frac{4\pi}{\lambda_d} n \sin^2 \left(\frac{\Theta}{2} \right) \tag{5.17}$$

Evaluating the cutoffs for the system parameters specified in Figure 5·2 gives values of $32\mu\text{m}^{-1}$ laterally and $10\mu\text{m}^{-1}$ axially. Note that the maximum axial frequency does not occur on the axial axis. In fact, on the axial axis the OTF is non-zero only at the origin. The lack of low-frequency axial support is known as the 'missing cone problem' due to the

Confocal Frequency Domain Support, n=1.51, NA=1.35
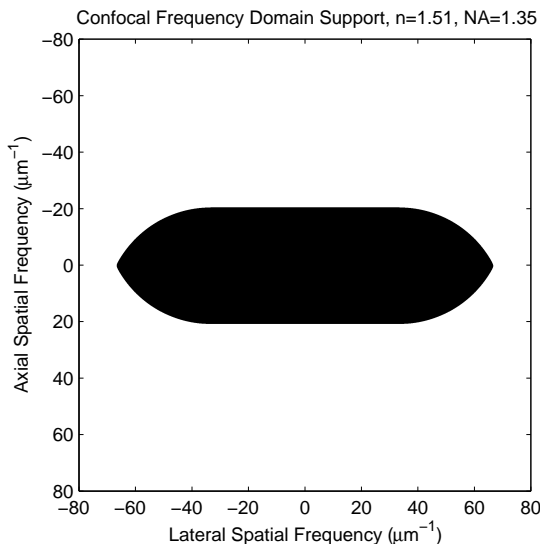
Figure 5·3: The support of the optical transfer function for a confocal microscope operating at a detection wavelength of 530nm and an excitation wavelength of 488nm. Note that this plot is a two-dimensional slice of the three-dimensional support that is cylindrically symmetric about the z axis.

shape of the low, predominantly-axial frequencies that fall outside the support. The result of the missing cone is that some low-frequency axial structures are not well resolved in a widefield microscope (Verveer et al., 1999).

5.3.2   Confocal Microscopy

The construction of a confocal microscope PSF was described in Chapter 4 with Equation 4.4, Equation 4.19 and Equation 4.22. For this system both the excitation and detection fields are focused. This means the confocal PSF is essentially the product of two widefield PSFs (one at the excitation wavelength and one at the detection wavelength). The confocal OTF (Gu and Sheppard, 1992; Sheppard and Gu, 1994) is thus the convolution of two widefield OTFs. In Figure 4·5 a confocal OTF is shown for a system using a numerical aperture of 1.35, an index of 1.51, circularly polarized excitation at 488nm and a detection wavelength of 530nm. The support is shown in Figure 5·3. Note that there is no longer a missing cone problem which is why confocal microscopy is often only regarded as a technique for improving axial resolution. However, the lateral spatial frequency cutoff

is also improved.

Physical descriptions of the operation of a confocal microscope often concentrate on the confocal detection pinhole as the means for rejecting light originating from outside the focal plane. This improves the axial resolution. While the confocal pinhole is important, it is also important to realize that the focused excitation is a key difference from widefield microscopy. If the pinhole is completely opened ($s(\mathbf{r}) = 1 \leftrightarrows S(\mathbf{k}) = \delta(\mathbf{k})$) then the confocal OTF becomes a delta function (for the detection) convolved with a focused intensity pattern (for the excitation) which is very similar to the widefield system except the roles of excitation and detection are reversed. Closing the pinhole to a very small size gives the OTF shown in Equation 4·5 with the support shown in Figure 5·3.

The Fourier-domain cutoffs can be calculated. The lateral cutoff is given below.

$$(\mathbf{k}_{c,\text{Conf.}})_x = (\mathbf{k}_{c,\text{Conf.}})_y = 4\pi \left(\frac{1}{\lambda_e} + \frac{1}{\lambda_d} \right) n \sin \Theta \tag{5.18}$$

The axial cutoff is as follows.

$$(\mathbf{k}_{c,\text{Conf.}})_z = 4\pi \left(\frac{1}{\lambda_e} + \frac{1}{\lambda_d} \right) n \sin^2 \left(\frac{\Theta}{2} \right) \tag{5.19}$$

These cutoffs correspond to values of $67\mu\text{m}^{-1}$ laterally and $21\mu\text{m}^{-1}$ axially in the example given.

5.3.3   4Pi Type C Microscopy

4Pi microscopy provides improved axial resolution by including an additional opposing objective lens. The Fourier-domain OTF construction (Gu and Sheppard, 1994) is then similar to the confocal case except that the amplitude transfer function is replaced by the sum of two opposing ATFs. Thus the building-block support changes from that of Figure 5·1 to that shown in Figure 5·4.

Examples of the resulting OTF were shown in the previous chapter in Figure 4·6 and the support is shown in Figure 5·5. This 4Pi system uses a numerical aperture of 1.35, an index of 1.51, circularly polarized excitation at 488nm and a detection wavelength of
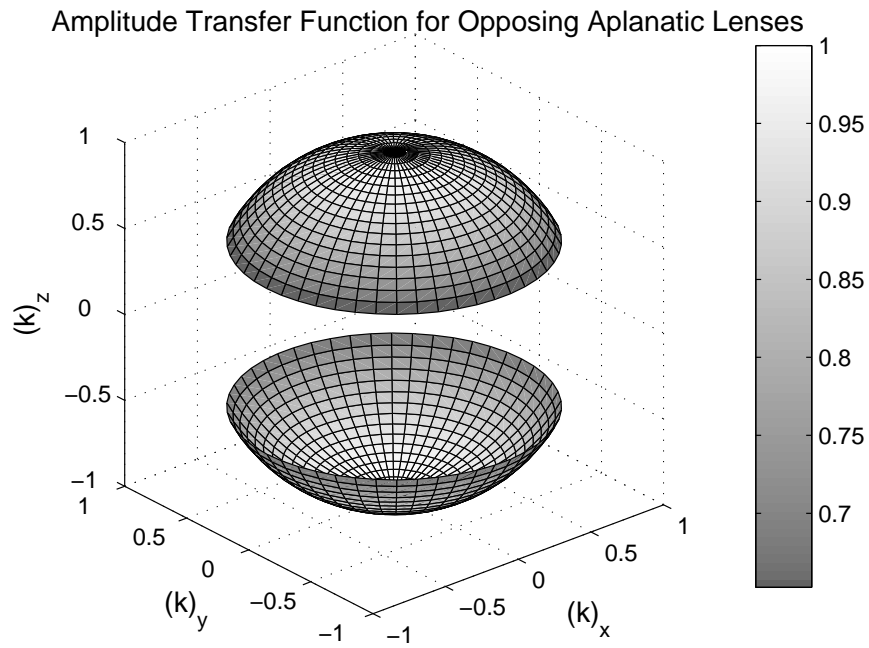
Figure 5·4: Graphical representation of the amplitude transfer function $\vec{A}(\mathbf{k}; \lambda)$ for a system using two opposing aplanatic lenses, each with a maximum collection angle of $\Theta = 65°$. The axes have been normalized.
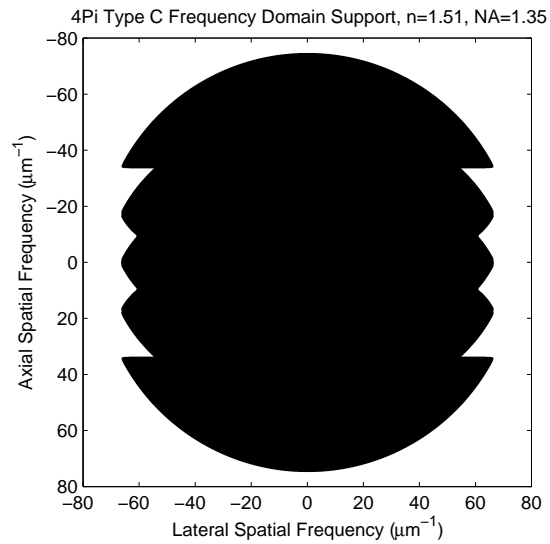
Figure 5·5: The support of the optical transfer function for a 4Pi Type C microscope operating at a detection wavelength of 530nm and an excitation wavelength of 488nm. Note that this plot is a two-dimensional slice of the three-dimensional support that is cylindrically symmetric about the z axis.

530nm. The cutoff frequencies can be calculated by the equations below.

$$(\mathbf{k}_{c,4\mathrm{Pi}})_x = (\mathbf{k}_{c,4\mathrm{Pi}})_y = 4\pi \left(\frac{1}{\lambda_e} + \frac{1}{\lambda_d} \right) n \sin \Theta \tag{5.20}$$

$$(\mathbf{k}_{c,4\mathrm{Pi}})_z = 4\pi \left(\frac{1}{\lambda_e} + \frac{1}{\lambda_d} \right) n \tag{5.21}$$

These cutoffs correspond to values of $67\mu\mathrm{m}^{-1}$ laterally and $75\mu\mathrm{m}^{-1}$ axially for the example system parameters.

5.3.4 SSFM Microscopy

Now that the origins of the OTF cutoff are known, the modifications that take a 4Pi to a SSFM system can be examined to see if they extend the cutoff. One of the critical aspects of a SSFM system is the delay between the two excitation/detection arms. These delays produce the characteristic spectral oscillations. However, when analyzed in terms of the OTF, the delay simply adds a constant phase to one of the two regions in Figure 5·4. This does not increase the support of the region shown and as a consequence, the support of the total OTF is not increased. Similarly, the possible offset between foci only produces a phase shift on one of the regions (as a spatial shift corresponds to a Fourier-domain phase modulation as stated earlier). Again, there is no increase in support.

The spectral detection does introduce a scaling to the support shown in Figure 5·4. Thus detection at a shorter wavelength expands this support and also expands the support of the overall OTF. The problem with this is that the range of detectable wavelengths is limited by the emission envelope of the fluorophore. This means that the resolution improvement is strictly limited by the minimum detectable wavelength and thus there is still no improvement over a 4Pi system operating at that same wavelength.

All these results are consistent with the OTFs shown in the last chapter. The detection wavelength was shown to change the size of the OTF but the effect of offset foci and path delays were simply to change the OTF within the same support. In short, the Fourier-domain construction of the SSFM OTF proves that the addition of path delay and foci

offset do not improve the resolution. Additionally, in terms of resolution, spectral detection does not provide anything more than the marginal benefit of detecting at the shortest wavelength available.

5.4 Pupil Filters

Adding phase and/or transmission properties to apertures has been proposed as a method for increasing the resolution achievable with a focused field (Toraldo di Francia, 1952) and is the subject of on-going research in the microscopy community (Sheppard and Hegedus, 1988; Sheppard, 1995; Grill and Stelzer, 1999; Neil et al., 2000; Martínez-Corral et al., 2002; Martínez-Corral et al., 2003; de Juana et al., 2003; Mugnai et al., 2003). These methods generally employ a spatial-domain analysis in a procedure often known as PSF engineering. These systems can also be examined using the Fourier-domain analysis described here. While this is typically not done in the PSF-engineering community, such an analysis leads to the conclusion that pupil filters are at best of limited use in resolution enhancement.

5.4.1 Effect on the OTF Support

A pattern across the lens pupil has a direct effect on the ATF. The spatial pattern (a function of x and y in the pupil plane) can be defined as a function of $\mathbf{k}$, as the lens transmission pattern is taken directly into the Fourier domain representation of the focused field according to the relation in (McCutchen, 1991). This transformation is easily derived geometrically.

$$\frac{(\mathbf{k})_x}{(\mathbf{k})_z} = \frac{x}{f}, \qquad \frac{(\mathbf{k})_y}{(\mathbf{k})_z} = \frac{y}{f}, \qquad |\mathbf{k}| = \frac{2\pi n}{\lambda} \tag{5.22}$$

Here f is the focal length of the lens. If the pupil filter patterns (one for each illumination polarization) are $R_x(\mathbf{k}), R_y(\mathbf{k})$ then the effect of the pupil filter on the ASFs is shown below.

$$\vec{A}_x(\mathbf{k};\lambda) \quad \overset{\text{Pupil Filter}}{\longrightarrow} \quad R_x(\mathbf{k})\vec{A}_x(\mathbf{k};\lambda) \tag{5.23}$$

$$\vec{A}_y(\mathbf{k};\lambda) \quad \overset{\text{Pupil Filter}}{\longrightarrow} \quad R_y(\mathbf{k})\vec{A}_y(\mathbf{k};\lambda) \tag{5.24}$$

Note that the lens patterns correspond to a physical variation in a lens that is being illuminated by a plane wave. This formulation is particularly useful as it can also model an arbitrary field distribution across the lens pupil. A field with x- and y-polarized components of $R_x(\mathbf{k})$ and $R_y(\mathbf{k})$ (where $\mathbf{k}$ is mapped to $\mathbf{r}$ by Equation 5.22) incident on the lens is equivalent to a lens with those pupil patterns being illuminated by a plane wave. Popular applications such as radially polarized beams (Youngworth and Brown, 2000; Neil et al., 2002) or STED depletion beams (Klar et al., 2001) can then be analyzed this way. Lens abberations can also be modeled in this framework — the pupil functions become the combination of Zernike polynomials that describes the abberation (Born and Wolf, 1999).

The Fourier domain effect of a pupil filter is shown in Equation 5.23 and Equation 5.24 and it can be seen that they do not increase the support of the ATF. Since the OTF support depends only on the ATF support for the systems considered, the pupil filters do not increase their support. Like SSFM, all that they do is modify the OTF within this support. In certain applications this may be useful, e.g. radial polarized beams have no lateral polarization component at the focal spot and can therefore be used for polarization selective measurements. While the effective pupil filter does not change the achievable resolution in this case, it does remove cross-talk from other polarization measurements at the focal spot.

The Fourier-domain support result will be demonstrated with an example. In (Martínez-Corral et al., 2003) a shaded-ring filter is used in confocal microscopy. The lens used has a numerical aperture of 1.2 and is used in a water-immersion system. This leads to a maximum collection angle of $\Theta = 64.5^o$. The pupil filter is applied to the excitation light at a wavelength of 350nm. A circularly polarized excitation beam will be used and detection will occur at 440nm. An illustration of the ATF for the shaded-ring filter is shown in Figure 5·6.

The resulting normalized point spread functions were calculated on a 20nm grid and are shown in Figure 5·7. As claimed in (Martínez-Corral et al., 2003), the pupil filter has altered the axial profile of the PSF in such a way that it appears more localized. Since a narrower
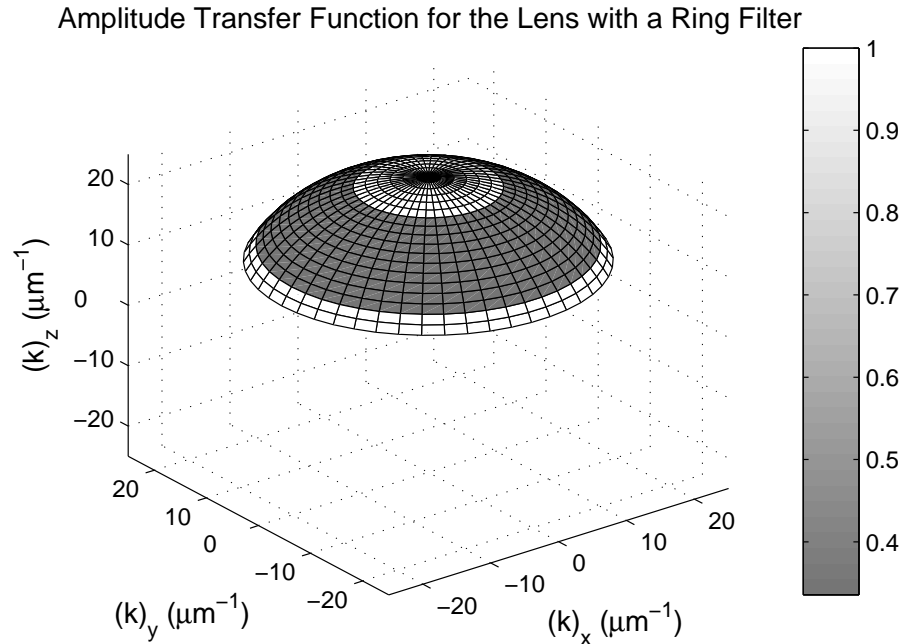
Figure 5·6: Graphical representation of the amplitude transfer function for the shaded-ring lens used in the pupil-filter example.

axial PSF is observed, it is argued that superresolution has been achieved. However, this analysis neglects any post-processing of the data and does not systematically examine the information passed by the microscope. These issues can be addressed further by looking at the OTFs of the relevant systems. These OTFs are shown in Figure 5·8.

Calculating the lateral and axial cutoff frequencies in this system gives $12.3\mu\mathrm{m}^{-1}$ and $3.88\mu\mathrm{m}^{-1}$ respectively. This agrees with the plots shown. The floor observed in the figures represents the effects of numerical calculation error in the simulations. Note that since the ASF is theoretically infinite in extent, truncation effects mean there will always be some level of calculation error in the OTFs. As expected, it can be seen that both OTFs observe the same cutoff frequencies. The pupil-filter does not move the cutoff but it does affect the OTF values within the support. The higher axial frequencies are passed more strongly in the pupil-filter system (by up to a factor of 3.5) which means they are more visible in the image. The heavier weightings on the edges of the support in Figure 5·6 are responsible for this relative increase in the OTF value toward higher frequencies.
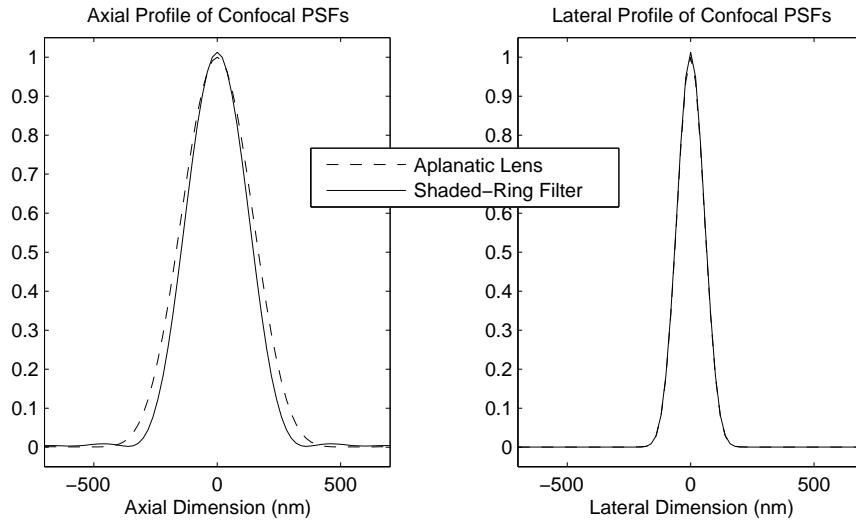
Figure 5·7: Lateral and axial profiles of the PSFs calculated in the pupil-filter example. One system uses aplanatic lenses while the other uses a shaded-ring pupil filter for the excitation. Note that the lateral profiles are too close to distinguish.
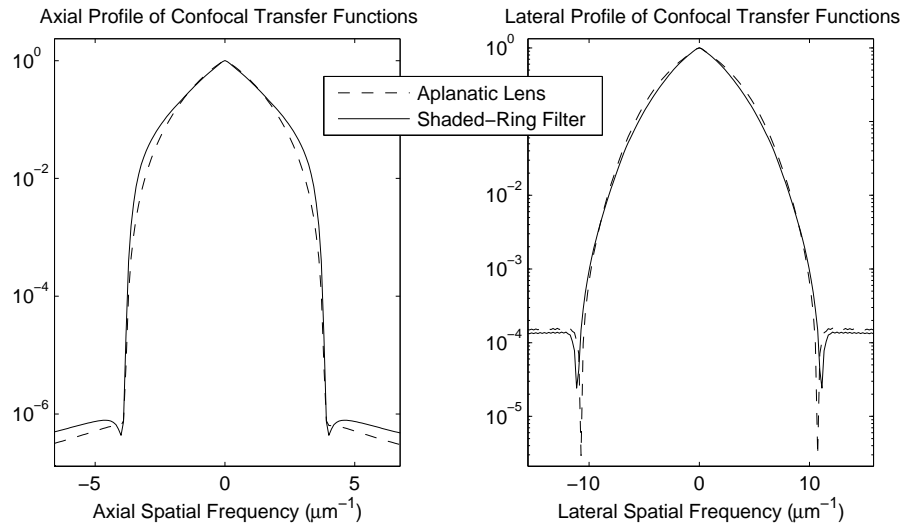
Figure 5·8: Lateral and axial profiles of the OTFs calculated in the pupil-filter example. One system uses aplanatic lenses while the other uses a shaded-ring pupil filter for the excitation.

Looking at the OTFs in Figure 5·8, it can be seen that the pupil-filter data could be synthesized from the aplanatic data by applying a high-pass filter in the noise-free case. High-pass filtering has the effect of enhancing edges in images (Parker, 1996) which can appear visually to be an improvement in resolution. However, as noted, the same effect could be synthesized using post-processing. Thus the pupil-filter can be regarded as a physical implementation of a signal-processing operation. If the cutoff frequencies were actually extended by some means, an edge-enhancing filter could not achieve the same effect as the spatial frequencies outside the smaller support are set to zero by observation through the instrument. This case would represent a true improvement in resolution using the new method. The difference between extending the OTF support and just raising the value of the OTF within it has been recognized in the literature (Cox and Sheppard, 1986) but is seldom addressed in the PSF-engineering-with-pupil-filters community.

5.4.2 Noise Considerations

It has been shown that the SSFM modifications or pupil filters cannot increase the support of the OTF. However, the OTF support may not be the only factor that influences the achievable resolution. While the cutoff frequencies do provide an upper bound on the resolution, regions within the support may be too noisy to reconstruct. If the OTF has a small but non-zero value at a certain value of $\mathbf{k}$ then the noise at that frequency may be higher than the data observed. Trying to reconstruct an estimate of the object at that $\mathbf{k}$ value will result in the noise-induced instability seen in Chapter 3. This means that noise can result in the achievable resolution being significantly less than the ideal noise-free resolution (Carrington et al., 1990; Stelzer, 1998). If pupil filters or similar modifications can increase the Signal to Noise Ratio (SNR) at such frequencies, they may increase the resolution achievable in the presence of noise.

In a Poisson noise model, the variance at a given data point is proportional to the expected count and independent of the noise at other points. This means that at a given point in the Discrete Fourier Transform (DFT) (Oppenheim and Schafer, 1989) of the data,

the noise variance is equal to the sum of the expected counts, which is in turn proportional to the OTF at $\mathbf{k} = 0$. The SNR (defined as the expected value of the signal divided by standard deviation) can then be calculated at each spatial frequency as shown below.

$$SNR(\mathbf{k}) \propto \frac{|H(\mathbf{k})|}{\sqrt{H(\mathbf{0})}} \tag{5.25}$$

The constant of proportionality above is determined by the ratio of the object's Fourier transform at $\mathbf{k} = 0$ (which influences the noise) to the object's Fourier transform at $\mathbf{k}$. Since this ratio is fixed for a given sample, the expression above will be used to calculate the SNR when comparing microscopes.

Another consideration that may be particularly important in fluorescence microscopy is the total illumination light to which the object is subjected. Photobleaching causes fluorophores to cease emitting after a certain amount of exposure and this is obviously detrimental to imaging quality. For this reason it is particularly important to collect as high a proportion of emitted photons as possible. For example, in the shaded-ring filter example the absorbing pattern is applied to the excitation lens, which lessens the illumination intensity and a clear detection lens is used. Applying the filter to the detection lens and using a clear excitation pupil would produce a similar OTF but would be much worse in terms of photobleaching. Swapping the filter in this manner would give higher excitation intensity and a lower detection efficiency which would exacerbate the photobleaching problem.

Photobleaching limits the amount of time for which an object can be effectively imaged (as fluorophore emission will decrease with time). Imaging time can also be limited by practical considerations such as object movement, instrument drift or imaging throughput requirements. The imaging time directly effects the scale of the PSF/OTF. The longer a sample is imaged, the more signal is collected and hence the imaging time can be incorporated into a scaling factor for the PSF/OTF. The scale of the OTF can also be determined by the total number of photons expected as the sum of the photons is determined by the

k = 0 values.

$$\text{Total Expected Photons} = D(\mathbf{0}) = H(\mathbf{0})O(\mathbf{0}) \tag{5.26}$$

From Equation 5.25 it can be seen that the SNR scales with the square root of the OTF scale. This mathematically describes how collecting more photons effects the SNR. While techniques such as 4Pi and confocal microscopy have extended OTFs, due to their scanning nature (which typically involves higher total illumination) they typically collect several times less photons than widefield microscopy (Pawley, 2000). So while a certain technique may extend the non-zero region of an OTF, low SNR in certain areas may make reconstruction of those frequencies impractical. For this reason it is important to consider the SNR of a system along with its OTF support when characterizing imaging quality. This is made clear by considering imaging systems with abberations. As mentioned earlier, abberations do not change the ATF support and therefore do not change the OTF support either. However even with sophisticated post-processing an aberrated system may not achieve the same imaging quality as a flaw-free system due to a lower SNR. The classic example of this is the Hubble space telescope where much effort was expended in mathematically correcting data from the flawed system (Hanisch et al., 1997). The result was still inferior to the images obtained when the telescope was physically repaired.

From Equation 5.25 it can be seen that a plot of the OTF also serves as a scaled plot of the SNR. When the OTFs in Figure 5·8 were calculated, the excitation intensities were normalized so that identical photobleaching would occur with either system. This means the OTFs shown give a SNR comparison for a constant photobleaching rate. From this viewpoint it can be seen that the shaded-ring filter does provide advantages in terms of the SNR. This means that more spatial-frequencies may be able to be reconstructed in a noisy environment — post-processing cannot provide this advantage. From this standpoint, pupil filters may improve imaging quality, in a noisy environment, if not the ultimate achievable resolution.

Another interesting feature is zero-crossings within the OTF support. While the maximum passable spatial frequencies can be calculated using the methodology outlined ear-
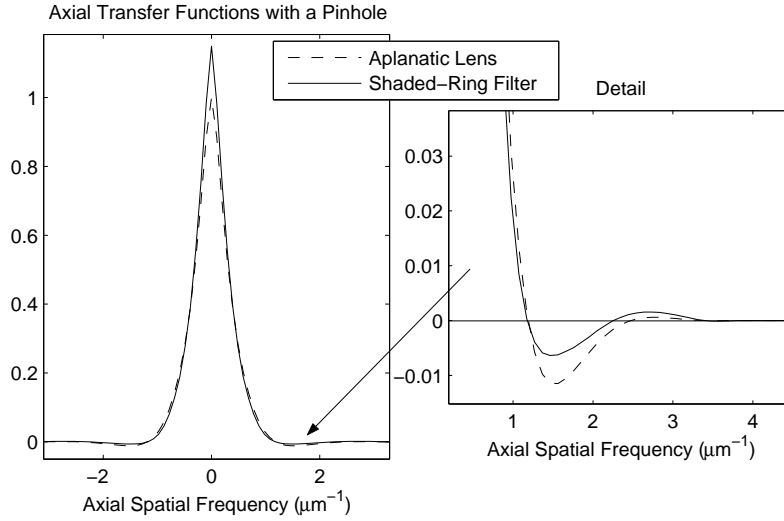
Figure 5·9: Axial profile of the OTFs calculated in the pupil-filter example when a large detector pinhole is used. One system uses aplanatic lenses while the other uses a shaded-ring pupil filter for the excitation.

lier, there may be zero-crossings within this range. These points lead to unobservable components in the data. Applying a pupil filter may move the zero-crossings as shown in Figure 5·9. In this example a large detector pinhole is included in the shaded-ring filter example. One of the effects of this pinhole is to induce zero crossings (Gu and Sheppard, 1992). The result seen leads to the possibility of using pupil-plane filters to deliberately place the zero-crossings in more convenient locations.

The OTF figures for SSFM in Figure 4·8 are normalized so that $H(\mathbf{0}) = 1$. This corresponds to data collection such that the same total number of photons can be expected when using any of the system parameters shown. Some of the more unattractive looking OTFs (e.g. the case where $\phi_e = 180°, \phi_d = 0$) have large gaps in the OTF but higher-than-standard values at the outer axial frequencies. This corresponds to a higher SNR in these regions but a very poor SNR in the low-frequency regions. Some potential use for this improved SNR is discussed in Chapter 7.

5.5   Summary

In Chapter 4 a method is given for calculating far-field microscopy PSFs from the amplitude spread function of a lens. In this chapter the same calculation methods are examined in the Fourier domain. Since the amplitude transfer function has a limited support, any PSF calculated by additions, scalings, products, spatial shifts, orientation changes, convolutions, conjugations and/or intensities of the ASF will result in an OTF with a limited support. These operations cover a wide range of far-field microscopy techniques including widefield, confocal, 4Pi and SSFM systems. The support of the OTF can be found by examining the Fourier domain calculation. This support defines an upper limit to the spatial frequencies passed by the system and correlates strongly to the highest resolution achievable. Using this approach, it can be seen that a SSFM system's support is no greater than that of a comparable 4Pi system operating at the lowest detection wavelength measured in the SSFM system. This gives a strong theoretical justification for the numerical results seen in the examples of Chapter 3 and Chapter 4 — that is that a SSFM system does not seem to outperform a 4Pi system.

A Fourier-domain analysis also allows one to see whether a given technique will increase the OTF support. If the ATF support is extended as in 4Pi microscopy (compare Figure 5·1 with Figure 5·4), then the OTF support may be extended. However techniques such as pupil filters, SSFM and others that don't change the ATF support will not extend the OTF support. This means the same spatial frequencies will still be lost during observation and hence the resolution will be limited in the same way. However, these techniques may alter the OTF values within the support which can have advantages in terms of signal-to-noise properties. This is particularly true if outer values of the OTF can be raised in comparison to the $\mathbf{k} = \mathbf{0}$ value since the noise variance is proportional to $H(\mathbf{0})$. The SNR is also proportional to the square root of the total expected counts and this can be increased by collecting photons for a longer period of time. Issues such as photobleaching will ultimately limit the effective imaging time.

Chapter 6

Relaxing Scanning Requirements with Multi-Channel Microscopy

It has been shown that SSFM does not hold promise for resolution greater than that of other interferometric techniques such as 4Pi microscopy. However, SSFM is a simpler system than 4Pi in terms of equipment set-up and alignment. The ability to localize single fluorophore points or layers with a relatively simple instrument is also useful in certain applications (Moiseev et al., 2004b; Moiseev et al., 2004c; Moiseev et al., 2006). Additionally, while 4Pi systems involve scanning the focal point through the object, SSFM can obtain relatively high resolution axial information without the use of scanning. This improves image acquisition time and reduces the mechanical complexity of the system. The axial range probed by a single SSFM measurement is limited by the depth of focus of the objective. Since the high-NA lenses used to produce a small lateral spot also give a short depth of focus, lateral resolution and axial range are traded off by means of the objective NA. Thus for high-lateral-resolution and/or extended axial range SSFM measurements, it will be necessary to move the focal spot (i.e. scan) to take multiple measurements along the axial dimension.

In this chapter the idea of using SSFM to reduce the scanning required to a level below that of a 4Pi instrument is explored. The methodology developed is also applicable to other multi-channel instruments and an example is shown for multi-detector confocal microscopy. The performance of the method is also analyzed in the presence of noise. This work was published in two articles (Davis et al., 2004a; Davis et al., 2005a). Since fewer points need to be scanned, the hope is that this method will allow a reduction in image acquisition

time. The need for fast imaging has been recognized in biological research (Petty, 2004) and approaches such as spinning-disk confocal microscopy (Petráň et al., 1968; Xiao et al., 1988) and Multifocal Multiphoton Microscopy (MMM) (Bewersdorf et al., 1998; Buist et al., 1998; Egner and Hell, 2000; Fittinghoff and Squier, 2000; Andresen et al., 2001) have been developed to produce fast, high-resolution imaging.

6.1   Sampling and Aliasing Background

Sampling involves representing a continuous signal with a discrete set of sample points from this function. In scanning microscopes this operation is done by physically collecting from one point and then moving a non-zero distance to collect the next point. The mathematics behind this operation are well known and the Nyquist condition (Jain, 1988) gives the minimum sampling rate for which the continuous signal can be reconstructed from the discrete samples. The Nyquist condition is well-known in the scanning microscopy community and is generally applied when choosing a sampling rate (Webb and Dorey, 1990). The Nyquist rate depends on the highest spatial frequency present in the function to be sampled. In this case the sampled function is the data which has non-zero frequency components only on the support of the OTF. Thus the maximum spatial frequencies (in x, y and z directions) are given by the vector $\mathbf{k}_c$ as calculated in the last chapter. The Nyquist sampling rates (in radians per unit space) are determined as follows.

$$\mathbf{k}_{\text{Nyq.}} = 2\mathbf{k}_c \tag{6.1}$$

This relation comes about because spatial sampling produces a replication of the Fourier spectrum. This idea is illustrated for one dimensional data in Figure 6·1. The replicated spectra are positioned at integer multiples of the sampling frequency. Satisfying the Nyquist criterion means that these spectra do not overlap and that the continuous signal can be found by low-pass filtering in order to remove the higher order spectral replicas. In the spatial domain, this low-pass filtering corresponds to interpolation. To satisfy the Nyquist criterion, the spatial sampling rates (given by the elements of $\mathbf{k}_{\text{Samp.}}$) must be above the

Continuous Signal's Fourier Spectrum

Sampled Signal's Fourier Spectrum
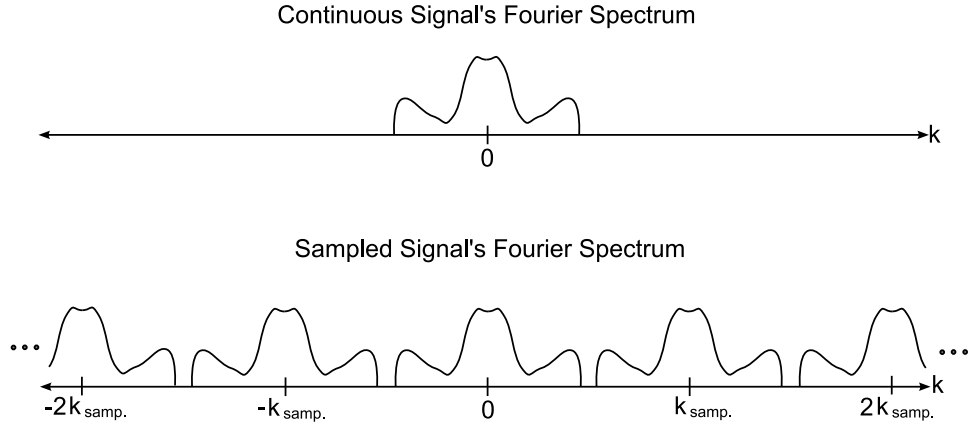
Figure 6·1: Illustration of the Fourier domain effect of sampling. When a continuous signal is sampled in space, it's Fourier transform is replicated at integer multiples of the sampling frequency.

Nyquist rate.

$$\mathbf{k}_{\text{Samp.}} \geq \mathbf{k}_{\text{Nyq.}} \tag{6.2}$$

If the sampling rate is not sufficient to ensure there is no overlap, the copies of the spectra add and a phenomenon known as aliasing occurs. In this case spatial frequencies from the baseband spectrum and higher order spectra overlap and cannot be separated by low-pass filtering. The result is that high spatial frequencies masquerade as lower spatial frequencies. As an example of this, consider a sinusoidal function sampled at a rate equal to its own spatial frequency. In this case all the sample values would be equal and the sinusoidal function would appear to be a constant. Another example of aliasing is shown in Figure 6·2 where a high frequency sinusoid and a low frequency sinusoid produce the same sample points. This indicates aliasing between these two frequencies. To distinguish the two a higher sampling rate would be required.

6.2   Unmixing Aliased Spatial Frequencies

Before considering the mathematics of removing aliasing effects with multi-channel measurements, the concept will be explained intuitively. An idealized dealiasing system is shown in Figure 6·3. In this system an object with two peaks is measured with two differ-
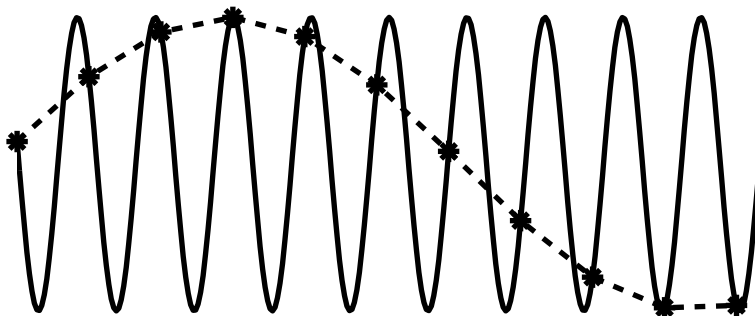
Figure 6·2: Illustration of aliasing. When the high-frequency solid curve is sampled at an insufficient rate the resulting sample points can also represent a low-frequency function (dashed curve).

ent systems — one which resolves the peaks and one which does not. The two data sets are then sampled at a sub-Nyquist rate which causes aliasing in the high-resolution data. This aliasing is clear since the resulting sampled high-resolution data has all points equal to zero. The dealiasing algorithm is then applied and it recognizes that the high-resolution data fixes the reconstruction value to zero at the sub-Nyquist sample points. The low-resolution data is non-zero which indicates that there are fluorophores present between the sample points. As illustrated, this allows a Nyquist-rate reconstruction to be found. The illustration is qualitative and represents an ideal situation but a general mathematical analysis follows.

While sampled data is a discrete set of numbers, it is convenient to analyze it as if it were a set of weighted delta functions placed at the sample locations. This treatment allows the spectral-replication effect of sampling to be easily shown — the sampled data is the continuous data multiplied by a delta function array, which leads to the continuous spectra convolved with a differently-spaced delta array in the Fourier domain. The result of this convolution is the replicated spectra. Sampled data will be denoted by an over-bar.

$$\bar{d}(\mathbf{r}) = \sum_{\mathbf{r}' \in \mathcal{S}} d(\mathbf{r}')\delta(\mathbf{r} - \mathbf{r}') \tag{6.3}$$
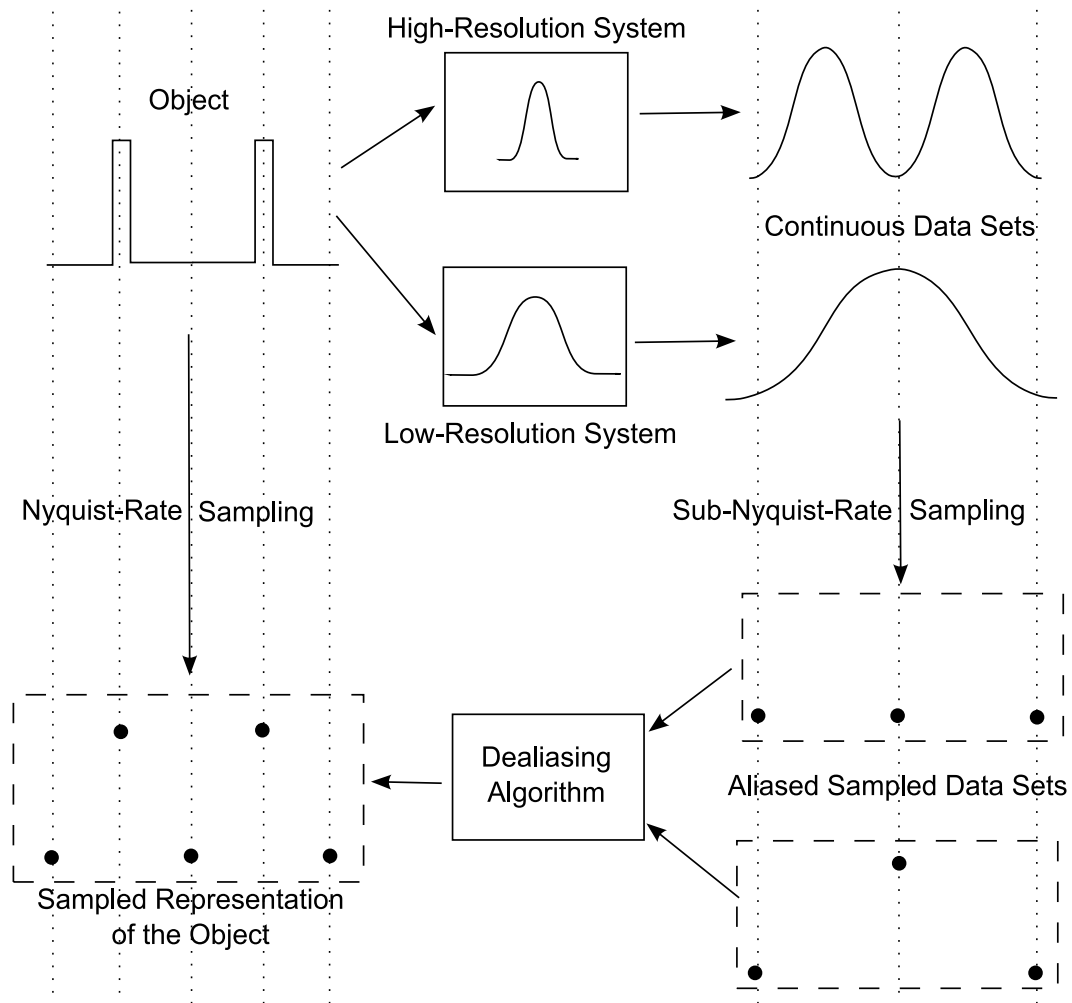
Figure 6·3: An idealized illustration of how the dealiasing procedure works. The object is measured through two different systems, one of which resolves the two peaks while the other does not. These data sets are sampled at a sub-Nyquist-rate and processed by a dealiasing algorithm. The result is the same as if the object were to be directly sampled at the Nyquist rate (as defined by the systems' maximum cutoff frequency).

The regularly-spaced sample points are contained in the set $\mathcal{S}$.

Under-sampling factors can be defined to quantify the amount by which the Nyquist criterion is violated.

$$N_x = \left\lceil \frac{(\mathbf{k}_{\text{Nyq.}})_x}{(\mathbf{k}_{\text{Samp.}})_x} \right\rceil \tag{6.4}$$

$$N_y = \left\lceil \frac{(\mathbf{k}_{\text{Nyq.}})_y}{(\mathbf{k}_{\text{Samp.}})_y} \right\rceil \tag{6.5}$$

$$N_z = \left\lceil \frac{(\mathbf{k}_{\text{Nyq.}})_z}{(\mathbf{k}_{\text{Samp.}})_z} \right\rceil \tag{6.6}$$

The maximum number of overlapping spectral points at a given $\mathbf{k}$ value is given by N (i.e. if N is greater than 1 aliasing may occur).

$$N = N_x N_y N_z \tag{6.7}$$

This means that N spatial frequency components mix into a single value in the aliased data. This is shown in the following expression where the sampled data's spectrum is written as a sum of different points in the continuous data's spectrum.

$$\begin{aligned} \bar{D}(\mathbf{k}) &\propto \sum_{u=1}^{N} D(\mathbf{k}^{(u)}) \\ &= \sum_{u=1}^{N} H(\mathbf{k}^{(u)}) O(\mathbf{k}^{(u)}) \end{aligned} \tag{6.8}$$

The N terms $\mathbf{k}^{(1)}, \mathbf{k}^{(2)}, \ldots, \mathbf{k}^{(N)}$ are the spatial frequencies in the object that alias into the spatial frequency $\mathbf{k}$ in the data. The final step in the above equation includes the observation model.

In Equation 6.8 the object's frequency components are mixed due to aliasing and cannot be recovered. However, multi-channel systems such as SSFM (each measured wavelength represents a single linear channel) provide additional information. The observation-aliasing

model of the mth channel of a M-channel system is as follows.

$$\bar{D}^{(m)}(\mathbf{k}) = \sum_{u=1}^{N} H^{(m)}(\mathbf{k}^{(u)})O(\mathbf{k}^{(u)}) \qquad (6.9)$$

This system of equations can be written in matrix form.

$$\bar{\mathbf{D}}(\mathbf{k}) = \Omega(\mathbf{k})\mathbf{O}(\mathbf{k}) \qquad (6.10)$$

The elements of the equation above are simply collections of the terms from Equation 6.9.

$$\bar{\mathbf{D}}(\mathbf{k}) = \begin{bmatrix} \bar{D}^{(1)}(\mathbf{k}) \\ \bar{D}^{(2)}(\mathbf{k}) \\ \vdots \\ \bar{D}^{(M)}(\mathbf{k}) \end{bmatrix} \qquad (M \times 1) \qquad (6.11)$$

$$\mathbf{O}(\mathbf{k}) = \begin{bmatrix} O(\mathbf{k}^{(1)}) \\ O(\mathbf{k}^{(2)}) \\ \vdots \\ O(\mathbf{k}^{(N)}) \end{bmatrix} \qquad (N \times 1) \qquad (6.12)$$

The matrix $\Omega(\mathbf{k})$ is $M \times N$ and contains the OTF values involved in the observation-aliasing equation.

$$[\Omega(\mathbf{k})]_{mu} = H^{(m)}(\mathbf{k}^{(u)}) \qquad (6.13)$$

The term shown above is the OTF of the mth channel evaluated at the uth aliased spatial frequency.

The form of Equation 6.10 immediately suggests a method for recovering the object, without aliasing, from the multi-channel data set. If $\Omega(\mathbf{k})$ can be inverted at all spatial frequencies of interest, then the object's spatial frequencies can be unmixed and appropriately arranged in Fourier-space to produce an estimate of the object. Sampling below the Nyquist rate in one-dimensional, noise-free systems was proposed some time ago (Papoulis, 1977) and more recent results have extended the theory to multi-dimensional, ill-posed and

noisy systems such as this one (Cheung and Marks II, 1985; Galatsanos et al., 1991; Cheung, 1993; Harikumar and Bresler, 1999). This idea of alias-free transmission through multiple sub-Nyquist channels has much in common with the digital communications concept of a filter bank (Vaidyanathan, 1992). It is also possible to change the sampling pattern to some benefit (Cheung and Marks II, 1990; Venkataramani and Bresler, 2003) but that approach is not considered here. The model of Equation 6.10 will be used to define a simple, intuitive inversion-dealiasing procedure that gives some intuition into the expected performance. Many standard approaches (including Tikhonov regularization, the Richardson-Lucy algorithm or the methods presented in the aforementioned multi-channel literature) could also be applied to this problem since a simple linear model has been formulated.

The inversion method used here will be based on the simple idea of inverting the matrix in Equation 6.10. If $\Omega(\mathbf{k})$ can be inverted at all $\mathbf{k}$ then the object could be reconstructed perfectly in a noise-free environment. This would require that there be at least as many observation channels (M) as the undersampling factor (N). However, even if $M \geq N$ the reconstruction process will be ill-posed due to nullspace and noise-amplification problems as discussed in Chapter 3. This necessitates including some sort of regularization when inverting $\Omega(\mathbf{k})$. This can be done by using the Truncated-Singular-Value-Decomposition (TSVD) inversion procedure (Hansen, 1998). Simply put, any $\mathbf{O}(\mathbf{k})$ component passed by $\Omega(\mathbf{k})$ with a gain less than the regularization parameter α will not be recovered by applying the TSVD inverse. The TSVD inverse will be denoted by $\Phi(\mathbf{k}; \alpha)$ which is an $N \times M$ matrix.

Applying $\Phi(\mathbf{k}; \alpha)$ to the data will result in estimates of the N components of the object's Fourier transform $O(\mathbf{k}^{(1)}), \ldots, O(\mathbf{k}^{(N)})$. The quality of these estimates, in regard to separating the N Fourier transform components, can be measured by the spectral purity measure given below.

$$\kappa(\mathbf{k}, \mathbf{k}^{(u)}; \alpha) = \frac{|[\Phi(\mathbf{k}; \alpha)\Omega(\mathbf{k})]_{uu}|}{\sqrt{\sum_{v=1}^{N} |[\Phi(\mathbf{k}; \alpha)\Omega(\mathbf{k})]_{uv}|^2}} \tag{6.14}$$

The aliasing effects of the undersampling are modeled in the matrix $\Omega(\mathbf{k})$ and the deconvolution processing by $\Phi(\mathbf{k}; \alpha)$. If the product of these two terms is the identity matrix then the deconvolution would be perfect. The spectral purity measure $\kappa(\mathbf{k}, \mathbf{k}^{(u)}; \alpha)$ measures how close this product is to the identity. The first spatial frequency argument ranges over the domain of the discrete Fourier transform of the data. The second argument takes on N values for each value of the first argument — these correspond to the frequency components in the object contributing to the data at $\mathbf{k}$. This means the second argument takes on values over the non-zero domain of the object. A purity measure of 1 means that no aliasing effects contaminate the estimate of the object at $\mathbf{k}^{(u)}$, while 0 means that the object at $\mathbf{k}^{(u)}$ is not reconstructed. A value between 0 and 1 indicates that aliasing effects are still present in the reconstruction at $\mathbf{k}^{(u)}$. In order to get a good reconstruction, it is desirable to have a well conditioned $\mathbf{H}(\mathbf{k})$ matrix (Golub and Van Loan, 1989) at as many spatial frequencies as possible. Qualitatively, this requires channels with significantly different transfer functions so that the rows in the matrix are not close to parallel (in a linear algebra sense).

6.2.1 SSFM Example

The methodology presented will be applied to a 4Pi/SSFM system. These results were published in (Davis et al., 2004a). As in Chapter 3, a test object will be defined, simulated data produced using a model of the system and then the image reconstruction algorithm will be applied. The test object used is shown in Figure 6·4. This is a spherically symmetric object that consists of layers with progressively closer spacing. This object is designed to have a broad spectrum so that the imaging quality is tested over the whole Fourier plane. The achievable resolution can be observed in the reconstruction by seeing which layer-spacings are resolved and which are not.

In the SSFM system simulated, a numerical aperture of 1.35 is used in a material with refractive index 1.51. The excitation light is circularly polarized, has a wavelength of 488nm and undergoes constructive interference ($\phi_e = 0$). Detection is done through a

Figure 6·4: The test object used in the SSFM simulations. An image of a lateral-axial slice of this spherically symmetric object is shown.

point pinhole and the detected wavelengths range from 500nm to 580nm. Three systems are considered — a single-channel instrument scanning at the Nyquist rate with a detection wavelength of 500nm and zero phase offset; a two-channel instrument scanning at half the Nyquist rate axially, detecting at 500nm and 580nm, and with detection phase offsets of $\phi_d(500\text{nm}) = 0$ and $\phi_d(580\text{nm}) = -120°$; and a five-channel instrument scanning at one quarter the Nyquist rate axially, detecting at 500nm, 527nm, 540nm, 553nm and 580nm, and with detection phase offsets of $\phi_d(500\text{nm}) = 0$, $\phi_d(524\text{nm}) = 90°$, $\phi_d(540\text{nm}) = 120°$, $\phi_d(557\text{nm}) = -120°$ and $\phi_d(580\text{nm}) = -90°$. The simulated data was generated on a 10nm grid before being resampled by each of the systems considered.

Reconstruction results for these systems are shown in Figure 6·5. The data used was noise-free and the regularization parameter set so that, in the Nyquist-rate case, the whole OTF support (see Figure 5·5) would be reconstructed. The Nyquist-rate reconstruction and the spatial frequencies used to construct it are shown in the figure. In the Nyquist-rate case there are no aliasing effects so the spectral purity takes on values of either 0 or 1. The
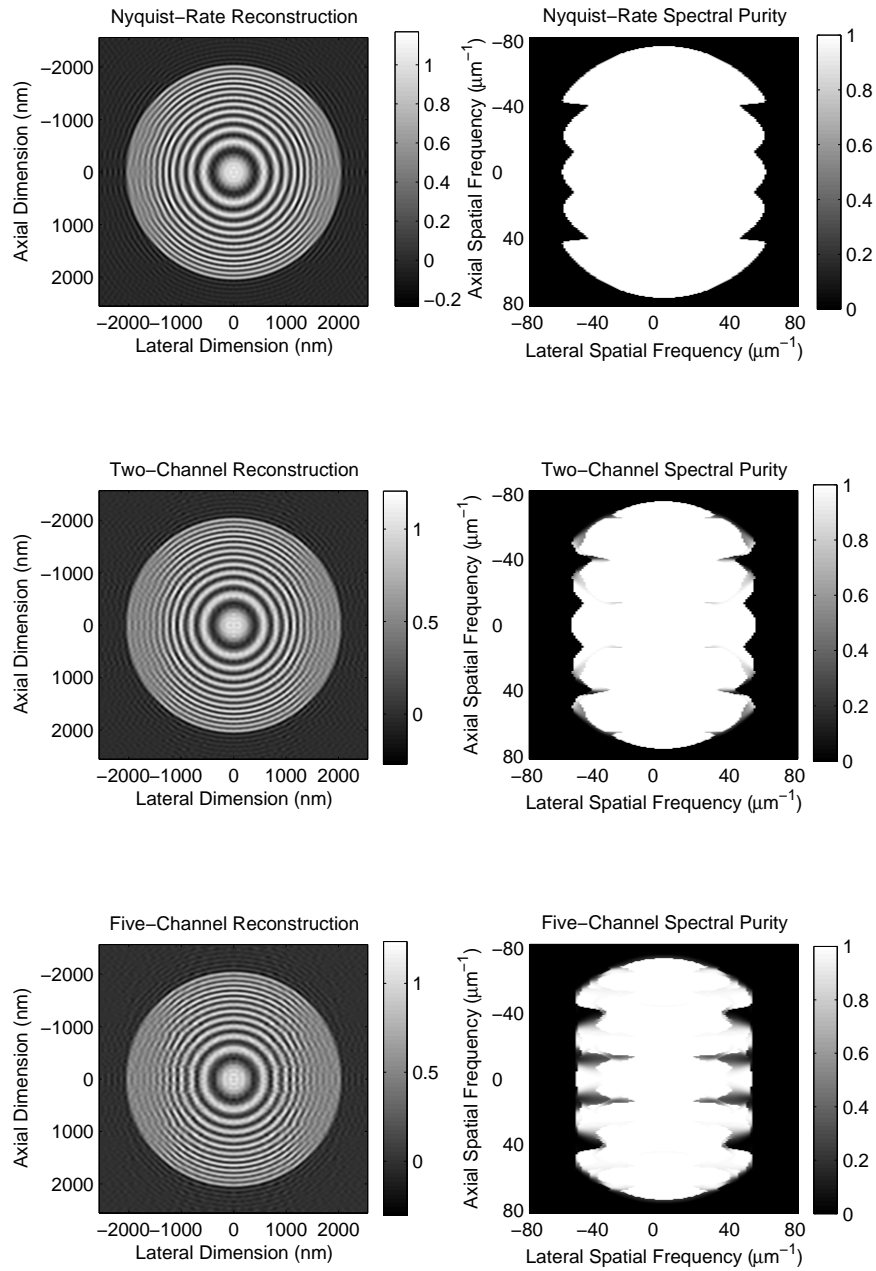
Figure 6·5: Lateral-axial slices of reconstructions of the SSFM test object from noise-free data. Reconstructions from a single Nyquist-rate channel, two half-Nyquist-rate channels and five quarter-Nyquist-rate channels are shown on the left. Corresponding spectral purity measures are shown on the right. Note that all sub-Nyquist sampling is done in the axial direction.

loss of the high spatial frequencies results in some ringing artifacts in the reconstruction (Sheppard and Wilson, 1981). Blurring on the lateral edges of the reconstruction is due to the fact that the lateral OTF support is too narrow to pass sufficient spatial-frequencies to allow these areas to be resolved.

The spectral purity of the two-channel reconstruction can be seen to be inferior to the Nyquist-rate case. This is due to aliasing effects which cannot be undone. The spectral purity now takes on values between 0 and 1 also. With close examination the reconstruction quality can be seen to be slightly worse but in general, the mixed spatial frequencies are well separated by the dealiasing processing. The data for the five-channel system was collected using only one quarter of the sample points used in the Nyquist-rate case. The detrimental effects of this are more visible in both the reconstruction and the calculated spectral purity. The structure of the object is still visible though and aliasing effects are not a significant problem.

The simulations shown here indicate that a SSFM system may operate below the Nyquist rate. This was suggested by the fact that axial localization of an object has been achieved using a system that does not scan (Swan et al., 2003). Unfortunately this interesting concept poses significant difficulties in implementation. The phase offsets ϕ_e and ϕ_d are sensitive to sub-wavelength optical path length changes in a physical path length that may be on the order of tens of centimeters. Maintaining the stability of a single phase in a 4Pi system is a challenging engineering problem which is significantly more complicated if multiple wavelengths are used. Dispersive elements in the optical path may produce a phase that varies rapidly with wavelength. This rapidly changing phase can cause the varying interference effects to wash out over the range of wavelengths within a spectrometer resolution unit. Even if dispersive elements could be compensated for, the object will introduce unknown dispersive elements, e.g. small changes in coverslip thickness will produce unknown dispersion effects. In the SSFM system of (Swan et al., 2003) the path length is precisely controlled using a spacer but as the numerical aperture of the lens increases, the spacer must become thinner to ensure that the reflected signal remains in

focus. This limits the range of phase offsets available and makes $\mathbf{H}(\mathbf{k})$ ill-posed (as the constituent transfer functions become increasingly similar). These practical difficulties (and general curiosity) encourage the multi-channel, reduced-scanning technique to be explored in the context of other systems. The next example presented is for a multi-detector confocal system and includes more detailed simulations including noise effects and reconstructions with significant aliasing present.

6.2.2 Multi-Detector Confocal Example

In a confocal fluorescence microscope, the light emitted by the fluorophores is focused on to the detector plane where much of it falls outside the detection pinhole. This discarded light represents a significant inefficiency in a method that is typically photon scarce. A practical method for collecting the light that falls outside the pinhole has been proposed using multiple detectors (Sheppard and Cogswell, 1990; Barth and Stelzer, 1994; Pawley et al., 1996). The example presented here discusses how a multi-detector confocal microscope may be used at a sub-Nyquist scanning rate by applying the theory developed in this chapter. This work has also been published elsewhere (Davis et al., 2005a).

In standard confocal microscopy a pinhole is used to reject light originating from out-of-focus planes in the sample and thus improve axial resolution. The smaller the pinhole, the more out-of-focus light is rejected but the signal level suffers as a result of the smaller detection aperture. The decreased SNR adversely affects image quality. The trade off is therefore between the ultimate achievable resolution and the SNR. Pinhole sizes are chosen to achieve a balance between these two ideas. In a multi-detector confocal microscope a relatively small central pinhole can be used to give high resolution capabilities, while detectors in the outer regions can be used to collect additional light that is usually discarded. The proposal here is that the extra detected light can be used to reduce scanning requirements.

The example system considered here uses a numerical aperture of 1.35 in an material with an index of refraction of 1.51. The excitation light has a wavelength of 488nm and is
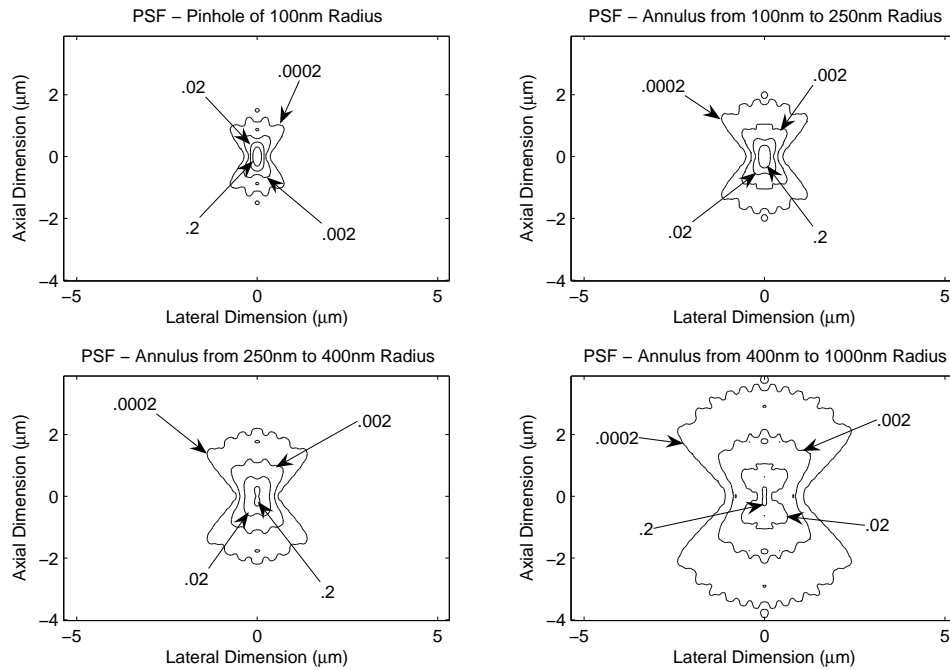
Figure 6·6: Contour plots of the confocal PSFs produced by a central detector and three concentric rings around that. The plots have been normalized so that the maximum value of the central pinhole's PSF is 1. Maximum values for the annuli (from the innermost to the outermost) are 1.12, 0.27 and 0.24.

circularly polarized. Detection is performed at a wavelength of 530nm. Four detectors are considered — one central circular pinhole with a demagnified radius of 100nm; an annulus with an inner radius of 100nm and an outer radius of 250nm; an annulus with an inner radius of 250nm and an outer radius of 400nm; and an annulus with an inner radius of 400nm and an outer radius of 1000nm. These represent a small pinhole surrounded by regions collecting progressively more out-of-focus light. The PSFs for these apertures are shown in Figure 6·6. It can be seen that the PSFs for the outer annuli collect significant contributions from outside the focal plane. The size of the PSFs also grows which indicates a decrease in resolution.

The OTFs of this multi-detector confocal system are shown in Figure 6·7. It can be seen that the OTF for the central pinhole has the broadest profile. This is to be expected
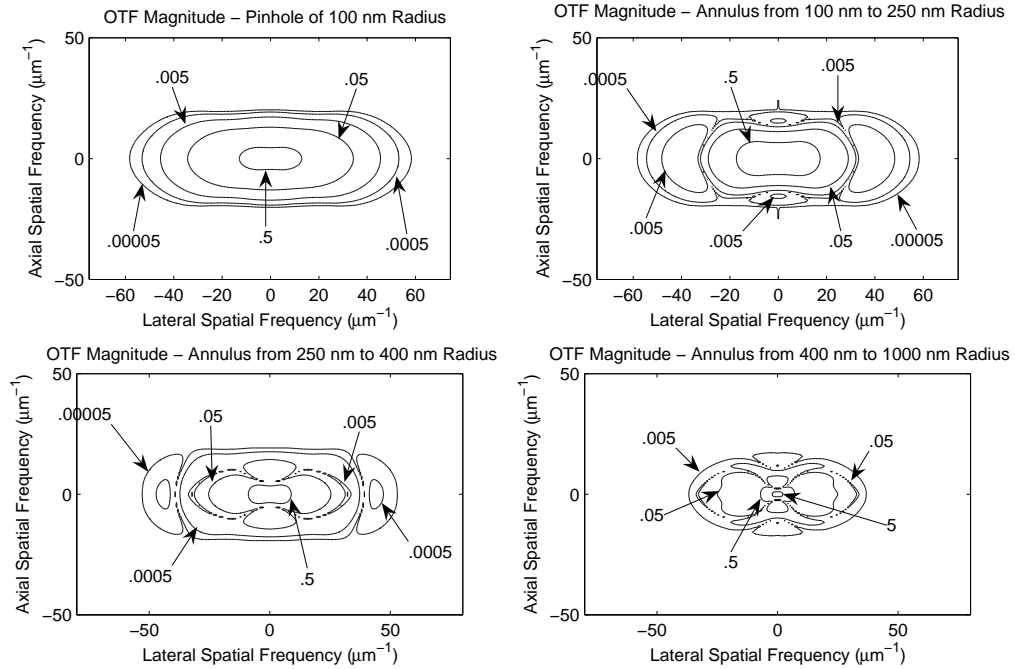
Figure 6·7: Contour plots of the confocal OTFs produced by a central detector and three concentric rings around that. The plots have been normalized so that the maximum value of the central pinhole's OTF is 1. Maximum values for the annuli (from the innermost to the outermost) are 3.11, 3.00 and 10.98.

as it should produce the highest resolution images. The annuli have progressively narrower support but do pass spatial frequencies near the origin strongly. This is also to be expected — the out-of-focus light contains a significant proportion of the photons but provides only low resolution.

The test object used in the multi-detector confocal simulations is shown in Figure 6·8. This is similar to the SSFM test object of Figure 6·4 except that it has fewer layers with more space between them. This is simply because a confocal system has lower resolution and can not be expected to resolve layers as closely spaced as the 4Pi system. The displayed object has a sample spacing of 20nm laterally and 62.5nm axially. The simulated data is also generated on this grid before being subsampled to the rate at which the examined system operates.
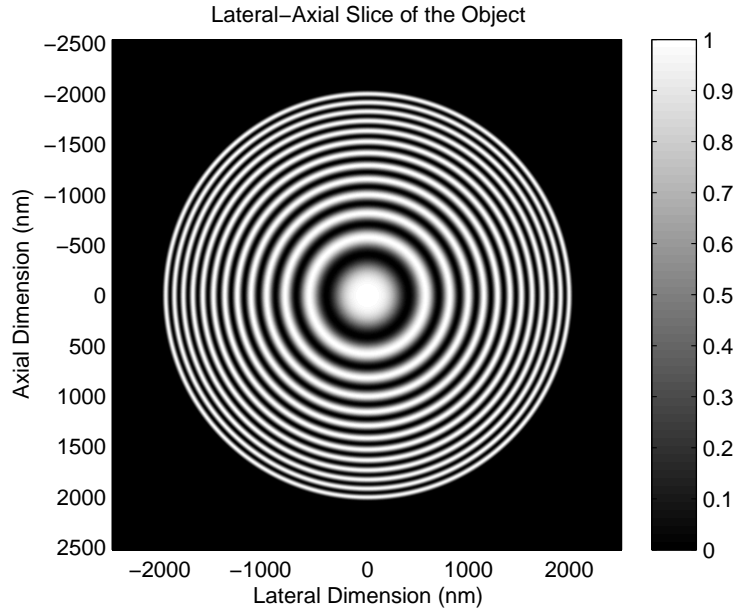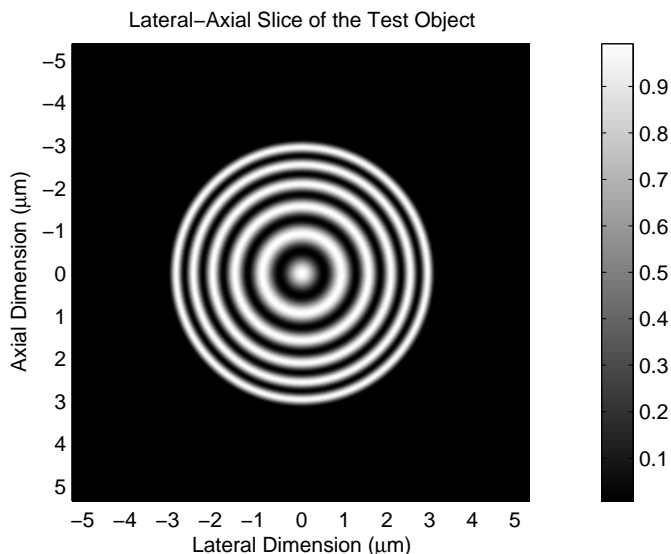
Figure 6·8: The test object used in the confocal simulations. An image of a lateral-axial slice of this spherically symmetric object is shown.

To illustrate the effects of aliasing, the object is sampled at different rates in Figure 6·9. These images are those that would be collected by two perfect imaging systems operating at different sampling rates. On the left the object is sampled at 40nm laterally and 125nm axially. The corresponding sampling rates are $157\mu\text{m}^{-1}$ and $50\mu\text{m}^{-1}$ while the theoretical cutoff frequencies for the confocal system (see Equation 5.18 and Equation 5.19) are $67\mu\text{m}^{-1}$ and $21\mu\text{m}^{-1}$. This means the sampling rate satisfies the Nyquist criterion of Equation 6.1 and the sample spacings of 40nm and 125nm will be used as the nominal Nyquist rates throughout these confocal simulations. This sampling rate also appears to capture the object well as seen in Figure 6·9. On the right of that figure the plot shows what happens when the object is sampled at half this rate axially — i.e. the axial sample spacing is 250nm. In this case the effects of aliasing are clear. The higher axial frequencies are not well captured and may even be interpreted as lower frequencies with a different orientation.

Noise-free reconstructions for the multi-detector confocal system are shown in Figure 6·10. Looking at Figure 6·7 it can be seen that the OTFs have very low values over a significant percentage of their support. This means that in a practical system it will prob-
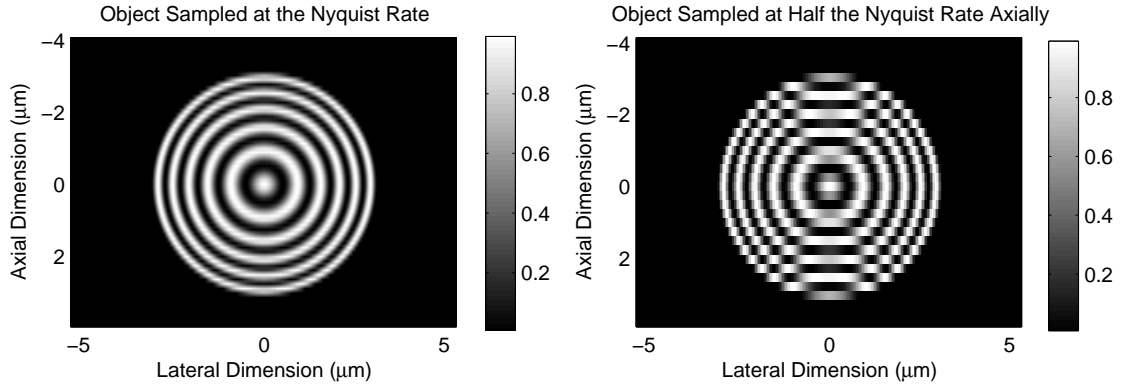
Figure 6·9: Aliasing illustrated by sampling the confocal test object above
the Nyquist rate (left) and below the Nyquist rate axially (right).

ably not be possible to reconstruct over the whole OTF support due to SNR concerns. A
more realistic regularization parameter will be used in this case and looking at the Nyquist
rate reconstruction's spectral purity measure it can be seen that a significant proportion of
the spatial frequencies within the OTF support are not reconstructed. The reconstruction
parameter α is high enough so that all the layers in the test object can be resolved, if only
barely for the high axial frequencies.

Since many of the higher spatial frequencies are not reconstructed, it could be argued
that the sampling rate could be lowered. To see what happens in this case a one-channel
system is considered operating at half the Nyquist rate axially. The spectral purity for
this system exhibits values between 0 and 1 (although this is not easily seen on the plot)
and has a reduced support. This indicates aliasing occurs and the number of retrievable
spatial frequencies is also affected. The effects of this are visible in the reconstruction
— the higher axial frequencies affected are poorly reconstructed when compared to the
Nyquist-rate case.

The first detector annulus is added to this system and the dealiasing processing applied.
The results are shown in the third plot of Figure 6·10. It can be seen that the spectral
purity is extended and the reconstruction is visibly improved. The result is still not as
good as the Nyquist-rate case but corresponds to collecting the data in half the time. The
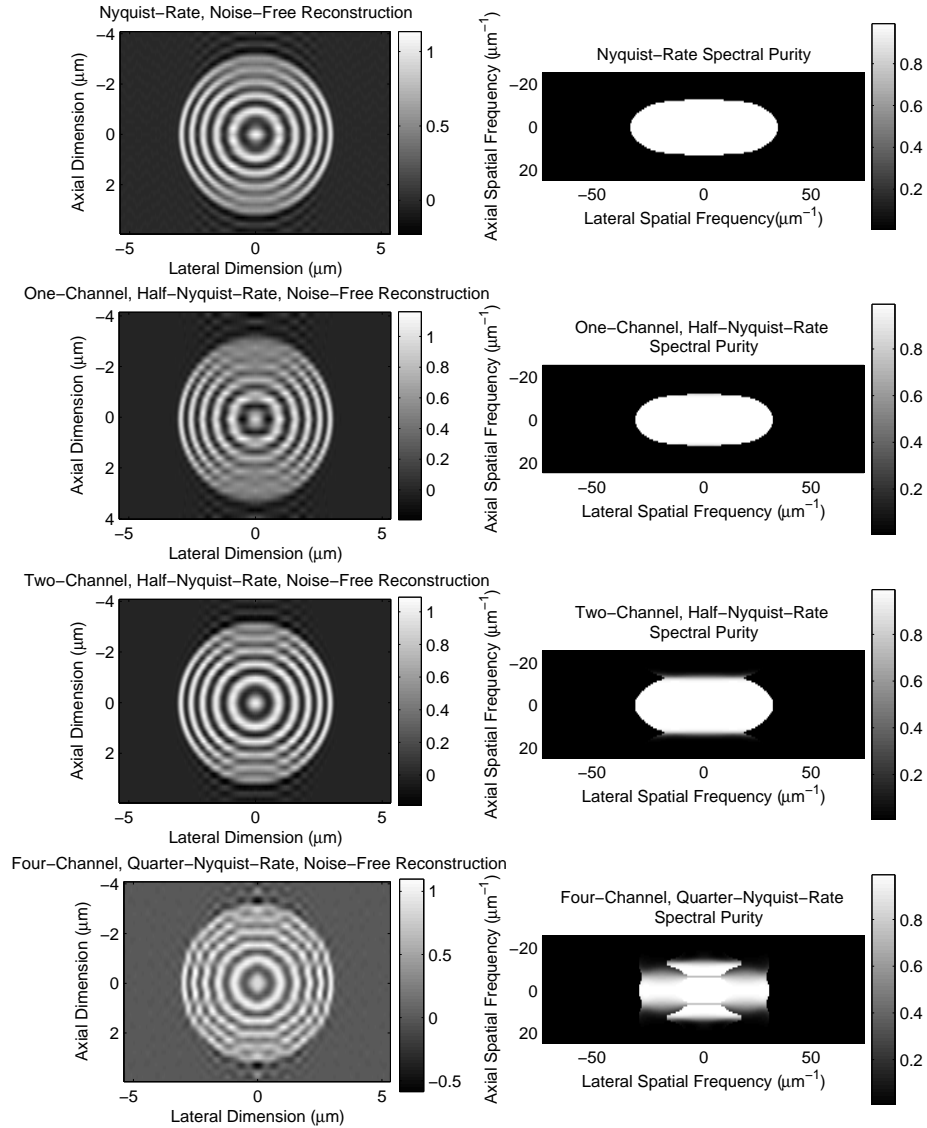
Figure 6·10: Lateral-axial slices of reconstructions from noise-free data and the corresponding spectral-purity measures. A single-channel Nyquist-rate reconstruction, a single-channel half-Nyquist-rate reconstruction, a two-channel half-Nyquist rate reconstruction and a four-channel quarter-Nyquist rate reconstruction are shown. When the Nyquist criterion is violated it is only in the axial direction. All reconstructions are for the same regularization parameter.
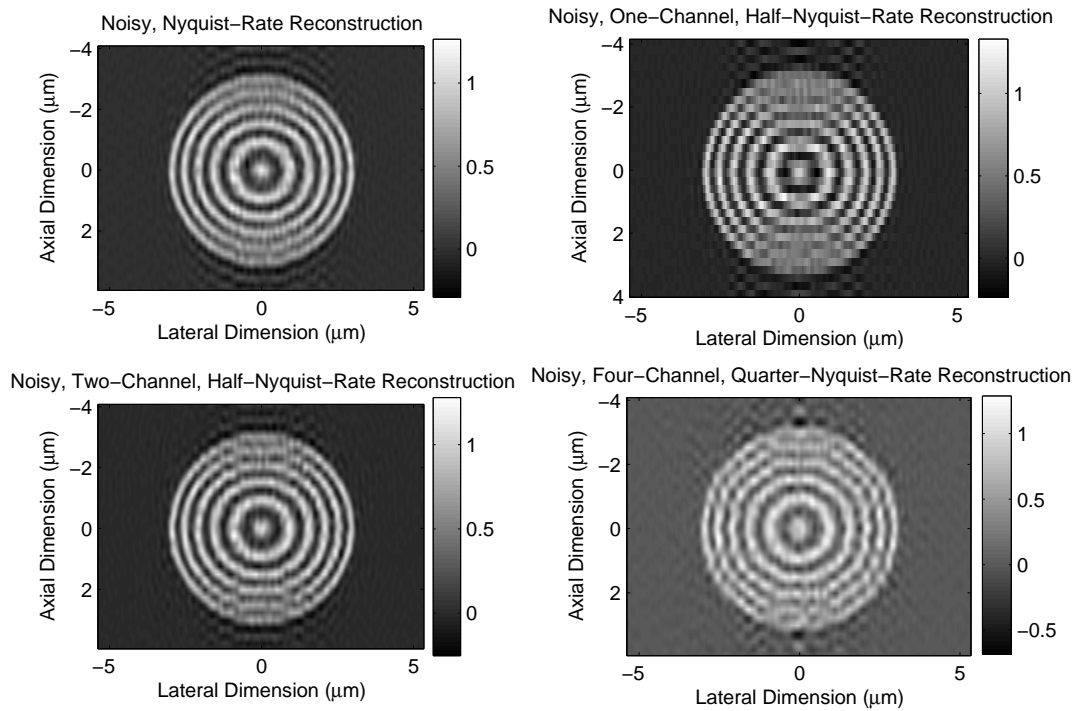
Figure 6·11: Lateral-axial slices of several reconstructions from noisy data. A single-channel Nyquist-rate reconstruction, a single-channel half-Nyquist-rate reconstruction, a two-channel half-Nyquist rate reconstruction and a four-channel quarter-Nyquist rate reconstruction are shown. All use the same regularization parameter.

key point in this reconstruction is that the higher axial frequencies are resolved.

The limits of this technique are further explored in the final plots of Figure 6·10. In this case the axial sampling rate is one quarter of the Nyquist rate. It can be seen that both the spectral purity and the reconstruction noticeably suffer. Some aliasing effects are visible in the reconstruction but only at the higher axial frequencies. Considering the axial sampling period is 500nm, this is not a bad result.

The reconstructions of Figure 6·10 are repeated in Figure 6·11 but with Poisson noise added to the data. A peak expected count of 500 in the central-pinhole data is used as the resulting SNR is consistent with the regularization parameter used in the noise-free data. This signal level may be somewhat high for a fluorescence instrument but is reasonable for
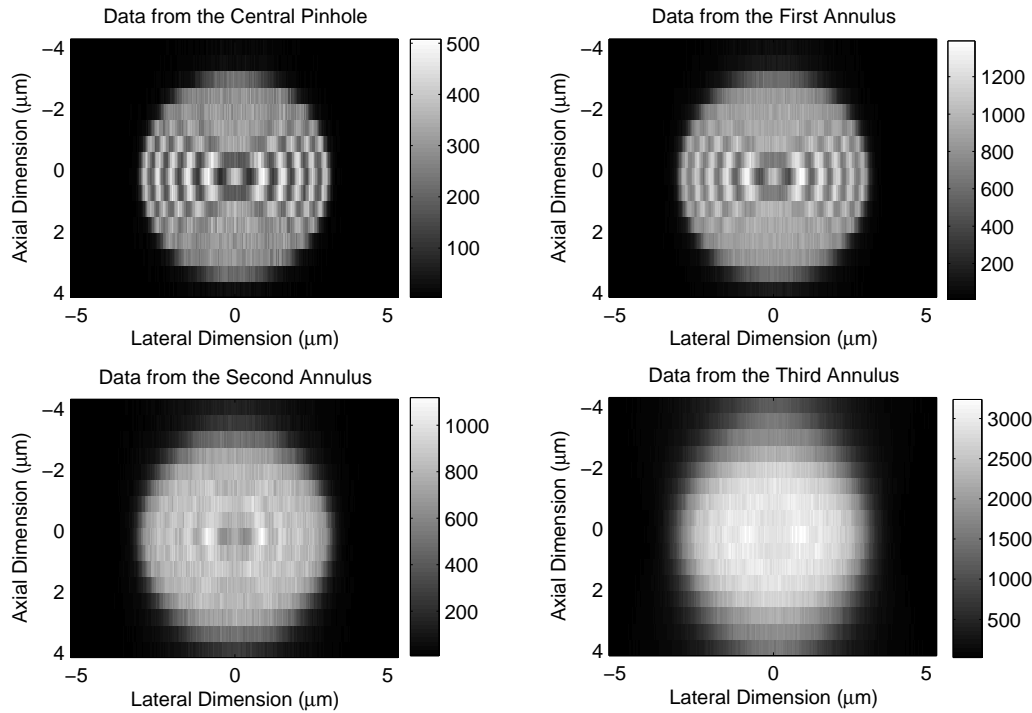
Figure 6·12: The four sets of simulated noisy data collected from the four-channel instrument operating at one quarter of the Nyquist rate axially.

other types of confocal systems, e.g. reflectance confocal microscopy. The reconstructions in Figure 6·11 do exhibit noticeable degradation due to noise, particularly in the high axial frequencies where high gains must be used to perform the reconstruction. However, the main message from these reconstructions is that the reconstruction process is stable in the presence of noise. This indicates that the regularization inherent in the TSVD does ensure stability with a correctly selected regularization parameter.

To emphasize the capabilities of the inversion-dealiasing procedure, the raw, noisy data for the quarter-Nyquist system is shown in Figure 6·12. Significant blurring and aliasing are visible in all of the data sets. With the algorithm developed, these data sets were processed into the bottom-left reconstruction of Figure 6·11.

6.3   Noise Analysis

So far it has been demonstrated that the principle of dealiasing can be applied to multi-channel microscopy systems that sample below the Nyquist rate. The important question that remains is whether or not this has any advantages over operating the same multi-channel microscope at the Nyquist rate. Reducing the sampling rate involves collecting less samples and thus reducing the image acquisition time. The image acquisition time could also be reduced by simply dwelling for less time at each sample point. Thus to evaluate the utility of the dealiasing idea one should compare a Nyquist-rate multi-channel microscope to the same multi-channel microscope collecting N times less points but dwelling for N times longer at each point. This ensures that both systems have the same image acquisition time and the resulting estimates of the object can be compared fairly. In order to perform such a comparison it is necessary to analyze the SNR in each instrument.

Taking a continuous Fourier transform of $\bar{d}(\mathbf{r})$ (Equation 6.3) plus a noise term gives the following relation.

$$\bar{D}(\mathbf{k}) = \sum_{\mathbf{r}\in\mathcal{S}}[d(\mathbf{r}) + w(\mathbf{r})]e^{-i\mathbf{k}\cdot\mathbf{r}} \tag{6.15}$$

This expression is simply the discrete-space Fourier transform. Note that the data term $d(\mathbf{r})$ is deterministic and randomness due to noise has been modeled by the term $w(\mathbf{r})$, which is $d(\mathbf{r})$ subtracted from a Poisson random variable with parameter $d(\mathbf{r})$. In Chapter 2 and Chapter 3 the discrete data was a random variable but here it is desirable to keep the effects of noise separate.

For convenience it will be assumed that the sub-Nyquist-rate system collects exactly $1/N$ times the samples of the Nyquist-rate system and dwells for N times as long at each. It is possible to compare systems with a non-integer factor between their rates but the analysis is somewhat more complicated (without adding insight). In both systems all M channels of data will be available. The Nyquist-rate system will be denoted by the subscript $_{\text{Nyq.}}$ and functions will be evaluated at the spatial frequencies $\mathbf{k}$ given by the Discrete Fourier Transform (DFT) (Oppenheim and Schafer, 1989), which can be implemented by the FFT.

Note that this discrete set of $\mathbf{k}$ values is sufficient to completely define the sampled data and the $\mathbf{k}$ values calculated lie between $-\mathbf{k}_{\text{Samp.}}/2$ and $\mathbf{k}_{\text{Samp.}}/2$.

$$
\begin{aligned}
\bar{D}_{\text{Nyq.}}^{(m)}(\mathbf{k}) &= \sum_{\mathbf{r} \in \mathcal{S}_{\text{Nyq.}}} [d^{(m)}(\mathbf{r}) + w_{\text{Nyq.}}(\mathbf{r})] e^{-i\mathbf{k}\cdot\mathbf{r}} \\
&= \sum_{\mathbf{r} \in \mathcal{S}_{\text{Nyq.}}} d^{(m)}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} + \sum_{\mathbf{r} \in \mathcal{S}_{\text{Nyq.}}} w_{\text{Nyq.}}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} \\
&= D_{\text{Nyq.}}^{(m)}(\mathbf{k}) + W_{\text{Nyq.}}(\mathbf{k})
\end{aligned}
\tag{6.16}
$$

Note that the deterministic component of the sampled data's spectrum $\bar{D}_{\text{Nyq.}}^{(m)}(\mathbf{k})$ is equal to the continuous spectrum $D_{\text{Nyq.}}^{(m)}(\mathbf{k})$ between $-\mathbf{k}_{\text{Samp.}}/2$ and $\mathbf{k}_{\text{Samp.}}/2$. This idea can be seen in Figure 6·1 and relies on the Nyquist criterion being met.

A similar quantification can be carried out for the sub-Nyquist-rate system (denoted by a subscript $_{\text{sub-Nyq.}}$). The dwell time is N times longer in this case so the signal from each point is proportionately stronger. The variance of each noise term is also N times higher according to the Poisson noise model.

$$
\bar{D}_{\text{sub-Nyq.}}^{(m)}(\mathbf{k}) = N \sum_{\mathbf{r} \in \mathcal{S}_{\text{sub-Nyq.}}} d^{(m)}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} + \sum_{\mathbf{r} \in \mathcal{S}_{\text{sub-Nyq.}}} w_{\text{sub-Nyq.}}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}
\tag{6.17}
$$

In this equation the DFT is taken over a different set of sample points than in the Nyquist-rate system. However the DFT can be rewritten over a common set of sample points. Summing N exponentials of frequencies $\mathbf{k}^{(u)}$ sampled on $\mathcal{S}_{\text{Nyq.}}$ results in an exponent of frequency $\mathbf{k}$ sampled on $\mathcal{S}_{\text{sub-Nyq.}}$. This result can be used in the following way.

$$
\begin{aligned}
\sum_{\mathbf{r} \in \mathcal{S}_{\text{sub-Nyq.}}} d^{(m)}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} &= \frac{1}{N} \sum_{u=1}^{N} \sum_{\mathbf{r} \in \mathcal{S}_{\text{Nyq.}}} d^{(m)}(\mathbf{r}) e^{-i\mathbf{k}^{(u)}\cdot\mathbf{r}} \\
&= \frac{1}{N} \left[\sum_{u=1}^{N} D^{(m)}(\mathbf{k}^{(u)}) \right]
\end{aligned}
\tag{6.18}
$$

This result is simply a derivation of the spectral replication idea shown in Figure 6·1 and

Equation 6.8. This result means Equation 6.17 becomes the following.

$$\bar{D}_{\text{sub-Nyq.}}^{(m)}(\mathbf{k}) = \sum_{u=1}^{N} D^{(m)}(\mathbf{k}^{(u)}) + W_{\text{sub-Nyq.}}(\mathbf{k}) \tag{6.19}$$

Comparing Equation 6.16 and Equation 6.19 forms the basis of the noise analysis. The variance of $W_{\text{sub-Nyq.}}(\mathbf{k})$ is equal to that of $W_{\text{Nyq.}}(\mathbf{k})$ because it is the unit-magnitude-coefficient sum of N less terms but each term has N times the variance.

The statement in Equation 6.19 is the same as that in the matrix relation of Equation 6.10 except that noise has been incorporated. A matrix equation can also be written for the Nyquist-rate system. Since there is no mixing between the frequencies ($N = 1$), the operator becomes a column matrix. Since the frequencies $\mathbf{k}^{(u)}$ cover the range of discrete frequencies between $-\mathbf{k}_{\text{Nyq.}}$ and $\mathbf{k}_{\text{Nyq.}}$, the Fourier-domain model for the Nyquist rate system can be written as follows. Noise has been neglected.

$$\bar{\mathbf{D}}_{\text{Nyq.}}(\mathbf{k}^{(u)}) = \begin{bmatrix} [\Omega(\mathbf{k})]_{1u} \\ [\Omega(\mathbf{k})]_{2u} \\ \vdots \\ [\Omega(\mathbf{k})_{Mu}] \end{bmatrix} O(\mathbf{k}^{(u)}) \tag{6.20}$$

This model uses the uth column of the sub-Nyquist operator $\Omega(\mathbf{k})$. Only one column appears since there is no mixing between the spatial frequencies. Applying the uth column of the inversion matrix $\Phi(\mathbf{k}; \alpha)$ to the Nyquist-rate data will result in an estimate of $O(\mathbf{k}^{(u)})$ that has the same noise variance as the sub-Nyquist case (since the noise terms have the same variance) and the same signal power but without any residual contamination from aliased spatial frequencies. Note that this analysis does assume the Fourier-domain noise terms are uncorrelated which is not guaranteed but should be a reasonable approximation for objects with many non-zero points. This signal-to-noise analysis indicates that a Nyquist-rate multi-channel system can do at least as well as a sub-Nyquist-rate multi-channel system with the same image acquisition time. This is not particularly surprising as both systems collect the same amount of photons. In fact the Nyquist-rate systems can

be expected to have potentially better performance as there is flexibility in processing the multiple-channel data unlike the sub-Nyquist system's data which must be processed in a manner to remove aliasing effects. Optimal processing for a multi-channel Nyquist-rate system is considered in the next chapter.

The consequences of this analysis for SSFM are simple. If lateral resolution and/or axial range requirements necessitate the need for an axially scanning SSFM system, then there is no advantage in performing this sampling below the Nyquist rate. Applications with an object that is confined axially or that a have low lateral-resolution requirement may result in a SSFM system that does not require scanning. In this case a SSFM system may have practical advantages over a complicated scanning instrument such as a 4Pi microscope.

6.4   Summary

Motivated by the scanning-free imaging of thin objects by a SSFM system, this chapter investigated how multi-channel microscopy could be used to relax scanning requirements. This analysis was also required due to the fact scanning may be needed in SSFM if extended axial objects are to be imaged or if high lateral-resolution requirements necessitate the use of a high-NA lens with a correspondingly shorter depth of focus. The hope was that if less sample points could be taken then the image acquisition time would be reduced.

A theoretical analysis showed how multi-channel systems could be used to allow sampling below the Nyquist rate. Simulations in a SSFM system and a multi-detector confocal instrument then validated this idea. While the effects of aliasing could be removed (with varying levels of success) an analysis of the signal-to-noise properties showed that the performance expected would not be better than a multi-channel Nyquist-rate system with the same image acquisition time. Essentially the dealiasing principle works but does not provide an advantage when applied to systems with a photon-counting noise model.

As mentioned in Section 1.3, SSFM has analogous properties to Fourier transform spectroscopy. Results from this field suggest that there may be potential advantages of SSFM over scanning instruments. A non-scanning SSFM system images fluorophores in

the vicinity of its depth of focus and provides axial information without scanning. The imaging of multiple spatial positions simultaneously is comparable to the simultaneous multi-wavelength detection achieved in Fourier transform spectroscopy. It has been shown that simultaneous acquisition of this type can lead to the 'multiplex advantage' (Fellgett, 1951; Fellgett, 1952). The multiplex advantage may provide significant gains in SNR, particularly when imaging a sparse object (Oliver and Pike, 1974). These results suggest that in addition to removing the need for mechanical scanning, SSFM may provide higher-quality images of axially confined, sparse objects. Evanescent excitation gives a means of assuring that the axial extent of the fluorescing region is limited.

Chapter 7

Suppressing Noise with Multi-Channel Microscopy

The previous section demonstrated how a multi-channel microscope can be used to reduce the number of sample points required in a scanning instrument. However it was also shown that this approach does not offer any improvements, in terms of image quality, over the same system operating at the Nyquist rate. However, it is still possible that a multi-channel microscope can offer improvements over a comparable single-channel instrument. This chapter explores this possibility and gives examples of microscopy systems that can be improved by modifying them to be multi-channel. The improvements are based on the idea that additional channels of the microscope can be used to collect photons that would otherwise be discarded. These photons are then used to improve the image quality using post-processing algorithms.

7.1 Image Reconstruction

Now that a Nyquist-rate system is being considered, the discrete data is once again easily related to the continuous data through sampling and interpolation. The observation equation returns to the familiar Fourier domain multiplication for linear shift-invariant systems.

$$D^{(m)}(\mathbf{k}) = H^{(m)}(\mathbf{k})O(\mathbf{k}) \tag{7.1}$$

This represents a linear relation between the object and the data and is thus sufficient for the implementation of standard reconstruction algorithms as applied in Chapter 3.

While it is possible to perform the inversion directly, the approach taken here will be to first perform a processing step that combines the multi-channel data into a single data
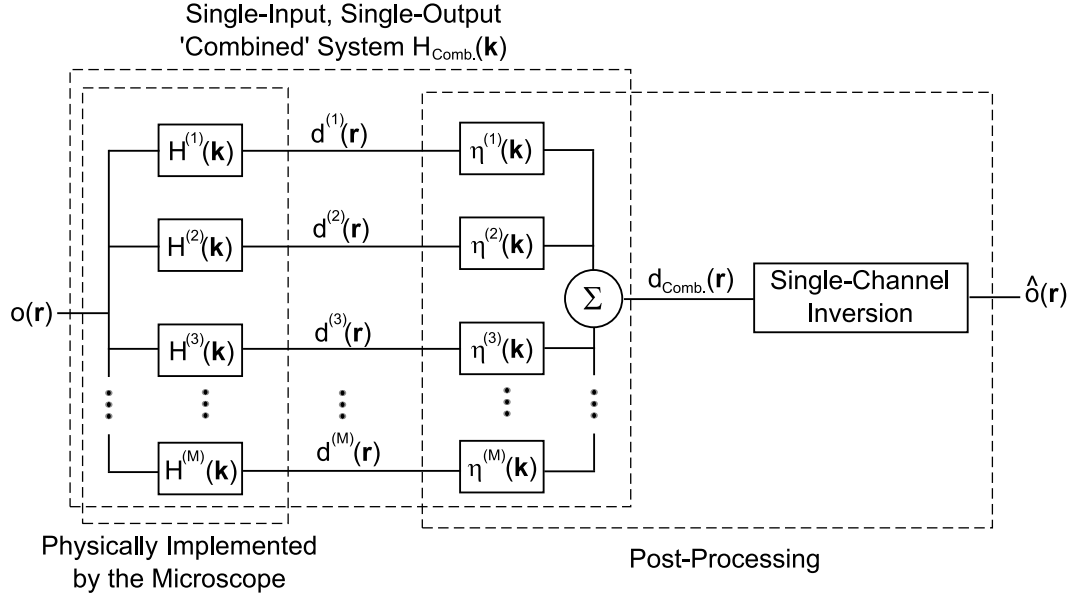
Figure 7·1: Systems diagram of the multi-channel microscope and the object estimation system. All systems are defined by their transfer function.

set similar to what is collected through a single-channel instrument. This approach allows standard single-channel inversion methods to be applied to the 'combined' data and also gives some intuitive results regarding the performance of the multi-channel system. The combination of the data will be carried out by an M-input, single-output linear shift-invariant system. This type of system can be described in the Fourier domain as shown below.

$$D_{\text{Comb.}}(\mathbf{k}) = \sum_{m=1}^{M} \eta^{(m)}(\mathbf{k}) D^{(m)}(\mathbf{k}) \tag{7.2}$$

Here $D_{\text{Comb.}}(\mathbf{k})$ is the combined data in the Fourier domain and $\eta^{(m)}(\mathbf{k})$ is the transfer function from the mth input to the output. A schematic of this observation and deconvolution system is shown in Figure 7·1.

Substituting the observation equation for the mth channel into the combination system of Equation 7.2 gives the following equation.

$$
\begin{aligned}
D_{\text{Comb.}}(\mathbf{k}) &= \sum_{m=1}^{M} \eta^{(m)}(\mathbf{k}) H^{(m)}(\mathbf{k}) O(\mathbf{k}) \\
&= H_{\text{Comb.}}(\mathbf{k}) O(\mathbf{k}) \tag{7.3}
\end{aligned}
$$

The term $H_{\text{Comb.}}(\mathbf{k})$ is the transfer function relating the combined data to the object and is defined as follows.

$$H_{\text{Comb.}}(\mathbf{k}) = \sum_{m=1}^{M} \eta^{(m)}(\mathbf{k}) H^{(m)}(\mathbf{k}) \tag{7.4}$$

This shows how the multi-channel microscope followed by the combination is mathematically similar to a single channel instrument with an OTF of $H_{\text{Comb.}}(\mathbf{k})$. The most significant distinction between a single-channel instrument and the combined multi-channel system is the statistics of the noise. The noise from a single physical observation is assumed to follow a Poisson model but even with Poisson inputs, the multi-channel combination system does not result in a Poisson distributed output. This is easily seen by noting that $d_{\text{Comb.}}(\mathbf{r})$ may have negative values. Another difference between a true single-channel instrument and the combined-channel system is that the combined data's noise is not uncorrelated between sample points. While the exact probability density function of the noise on the combined data is complicated, the variance at each $\mathbf{k}$ value is easily calculated.

The combination operation is defined by the functions $\eta^{(m)}(\mathbf{k})$. Specific values of these transfer functions will be defined by maximizing the SNR of the combined data $D_{\text{Comb.}}(\mathbf{k})$ at every $\mathbf{k}$ value. This is equivalent to maximizing the signal power with the noise power set. This is posed mathematically below.

$$\text{Maximize} \quad \sum_{m=1}^{M} \eta^{(m)}(\mathbf{k}) H^{(m)}(\mathbf{k}) O(\mathbf{k}) \quad \text{over} \quad \eta^{(m)}(\mathbf{k})$$
$$\text{subject to} \ \sum_{m=1}^{M} |\eta^{(m)}(\mathbf{k})|^2 H^{(m)}(\mathbf{0}) O(\mathbf{0}) = 1 \tag{7.5}$$

The maximization problem is easily solved using Lagrange multipliers to give the following result.

$$\eta^{(m)}(\mathbf{k}) \propto \frac{\left[H^{(m)}(\mathbf{k})\right]^{*}}{H^{(m)}(\mathbf{0})} \tag{7.6}$$

Note that the constant of proportionality is not important as it does not affect the SNR. The result above is intuitive, the weighting of the mth channel at a given spatial frequency is proportional to the strength with which that spatial frequency is passed and is inversely proportional to the noise on the channel (which is dictated by the DC component of the

data). In Equation 7.6 it can be seen that the combination filters are only zero if the corresponding data channel does not pass that particular $\mathbf{k}$ value. This means that all of the observed spatial frequencies reach the single-channel inversion block in Figure 7·1. The sort of Fourier-domain averaging seen here has also been used in other multi-channel systems (Gustafsson, 2000) and the idea of maximizing the SNR when designing the instrument has been employed before (Gauderon and Sheppard, 1999) but this did not consider a frequency-varying definition.

Following the application of this combination procedure, a single-channel deconvolution method can be applied. Note that not all algorithms would be appropriate, e.g. the Richardson Lucy algorithm is matched to Poisson noise and may not function well on the non-Poisson statistics of $d_{\mathrm{Comb.}}(\mathbf{k})$. Also, the criteria of Equation 7.5 defines a linear Minimum Mean Square Error (MMSE) combination under a stationary-noise approximation (Poisson noise is not stationary). This ensures that for a given single-channel-inversion system, the resulting image also has a minimal mean square error, assuming stationary noise. The remainder of this chapter contains simulation examples of the multi-channel inversion approach described here and also of the Richardson-Lucy algorithm applied directly to the multi-channel data. The multi-channel Richardson-Lucy algorithm is matched to Poisson noise, so these simulations give a point of comparison to the deconvolution approach developed above which assumes stationary noise.

7.2 Multi-Channel SSFM

Returning to the SSFM instrument, in Chapter 5 it was shown that the features of SSFM could not extent the OTF support of a 4Pi system. It was also noted that SSFM could change the OTF values within this support. The SSFM examples of Chapter 4 show that some system parameters (e.g. the $\phi_e = 0, \phi_d = 180°$ case in Figure 4·8) could increase the OTF values at high axial frequencies in relation to the $\mathbf{k} = 0$ value. These systems tended to have regions of very low OTF values in other areas though. This meant the improved SNR at certain regions was achieved at the expense of the SNR in other regions.

In short, different SSFM channels can have different SNRs at different spatial frequencies. The multi-channel combination algorithm provides a way of picking and choosing which channel is used at which spatial frequency — at any given spatial frequency the channel with the highest SNR is weighted the most heavily.

The SNR, as a function of spatial frequency, will be calculated for several 4Pi/SSFM systems. These systems all use objective lenses with a numerical aperture of 1.35 in a medium with an index of refraction of 1.51. Circularly polarized excitation light at 488nm with no phase offset ($\phi_e = 0$) is used along with a point detection pinhole. The calculations are done on a spatial grid with a spacing of 31.25nm in all three dimensions. Four systems are considered — a 4Pi system detecting wavelengths between 540nm and 560nm; a 4Pi system with a broad-detection filter that collects all wavelengths between 500nm and 600nm; a SSFM system collecting five equal-width channels in the 500-600nm range and with $\phi_d = 0$ for all channels; a SSFM system collecting five equal-width channels in the 500-600nm range with phase offsets of $\phi_d(510\text{nm}) = 0, \phi_d(530\text{nm}) = 135°, \phi_d(550\text{nm}) = 270°, \phi_d(570\text{nm}) = 45°$ and $\phi_d(590\text{nm}) = 180°$. These systems were modeled using OTFs calculated at wavelengths of 510nm, 530nm, 550nm, 570nm and 590nm. The OTFs were also scaled according to the envelope seen in Figure 3·2. This models the fact that the fluorophores do not emit uniformly across the whole detected spectrum.

The OTFs (or combined OTFs) for three of these systems are shown in Figure 7·2. The broad-detection 4Pi system and the $\phi_d = 0$ SSFM system produced results so similar that the two could not be distinguished on the plot, so only the former was plotted. The single-channel-4Pi OTF was normalized so that it had a $\mathbf{k} = 0$ value of 1. The other OTFs were scaled so that they produced the same noise variance in the data. This means the plots of the OTFs can also be regarded as plots of the SNR due to the form of the SNR definition (Equation 5.25) and the because all the OTFs are real and non-negative.

The fact that the broad-detection 4Pi system and the $\phi_d = 0$ SSFM system produced very similar results indicates that spectral detection and multi-channel processing is of little advantage in this case. This is because the OTFs vary very little if $\phi_e = \phi_d = 0$
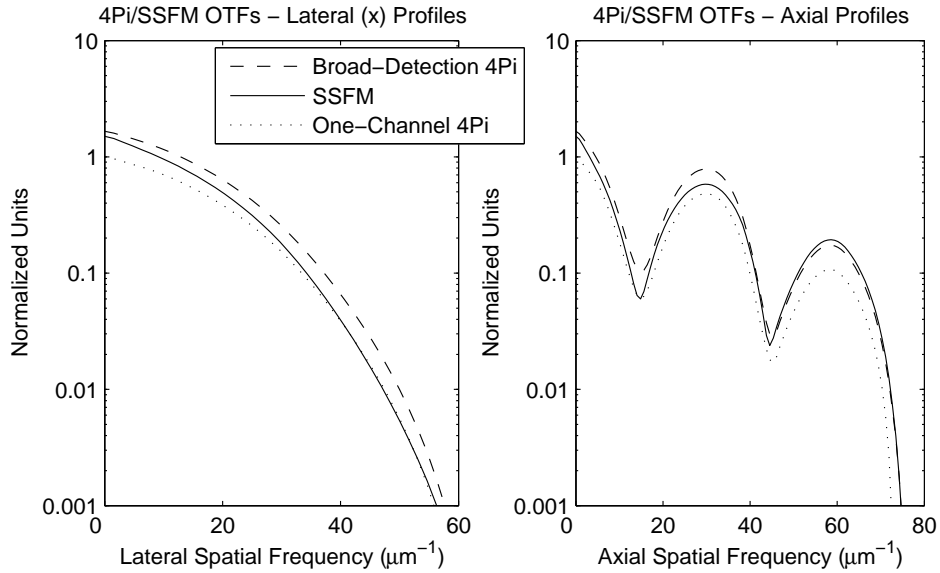
Figure 7·2: Lateral and axial profiles of the normalized OTF magnitudes for various SSFM/4Pi systems. A single-wavelength system, operating at 550nm is considered along with two multi-wavelength systems. The 'Broad-Detection 4Pi' employs a detector sensitive to all wavelengths in the 500-600nm band while the SSFM system has an excitation phase of zero and detection phase that varies with wavelength.

and only the wavelength is varied. Physically the broad detection corresponds to summing all the OTFs and Equation 7.6 also results in a simple unweighted sum if the channel OTFs vary only by a scale factor. This explains why the two OTFs considered here are so similar. It can be seen that the broad-detection 4Pi system does have a better SNR than the single-channel instrument — this is due simply to the increased number of photons collected. In practice a constant phase will be hard to maintain across the whole spectrum and will result in inferior results if broad detection is used.

The SSFM system does worse than the broad-detection 4Pi for most of the spectrum. Since the phase offsets produce detrimental effects for most spatial frequencies (see Figure 4·8), this is to be expected. However, as noted earlier, an improved SNR is noted at some of the higher axial frequencies. This is exploited by the multi-channel processing and does indeed produce a slightly higher SNR at certain frequencies, as can be seen in Figure 7·2. This advantage is very small and it would be difficult to justify the added

system complexity to realize it.

7.3   Multi-Channel Confocal Microscopy

As noted in Section 6.2.2, the pinhole in confocal microscopy allows an increased axial resolution but also results in discarded photons and hence poorer SNR. The tradeoff between signal level (SNR) and axial resolution through the variation of the pinhole size is well known (Sheppard et al., 1992). The topic of selecting an optimal pinhole size has also been addressed (Wilson, 1990; Gauderon and Sheppard, 1999). If the usually-discarded photons can be collected in additional channels there is the possibility of usefully incorporating them into the reconstruction of the object and hopefully improving the SNR without compromising axial resolution. This idea has been explored in the literature.

In (Sheppard and Cogswell, 1990) a confocal instrument with two-channels is considered. One channel collects a standard confocal signal through a small pinhole while the second channel collects from a larger pinhole. These two signals are weighted and then subtracted to produce improved imaging. This idea is further explored in (Heintzmann et al., 2003) where noise considerations are addressed more thoroughly. It is shown that postprocessing can synthesize a similar effect to the image subtraction but with better SNR. A Fourier-domain weighting and averaging of the channels (as proposed in this work) is shown to give superior results but it does rely on the assumed OTFs being accurate. It is also noted that it is important to simultaneously collect the two channels without unnecessarily discarding photons. A flexible instrument capable of such simultaneous collection is described in (Verveer et al., 1998) and alternative methods for processing the resulting data are given in (Verveer and Jovin, 1998).

The estimation of a spatially limited two-dimensional object from its diffraction-limited image is explored in a series of papers (Bertero and Pike, 1982; Bertero et al., 1982; Bertero et al., 1984a; Bertero et al., 1984b). A singular-value approach is employed to show that the small object allows a greater reconstruction resolution. It is noted that the focused illumination of a confocal microscope essentially limits the extent of the object imaged at

any one point and therefore recording the intensity over the whole detector plane (rather than just over the pinhole) holds the promise of improving the lateral resolution. A physical demonstration of the concepts developed in this series of papers is given in (Walker, 1983). A similar singular-value analysis is applied directly to confocal systems in another series of papers (Bertero et al., 1987; Bertero et al., 1989; Bertero et al., 1991; Bertero et al., 1992; Bertero et al., 1994). Here the axial resolution is considered and the mathematical processing implemented by means of transmission and phase (in the coherent instrument) masks in the focal plane. These masks are a physical implementation of a mathematical operation (similar to pupil filters but in the spatial domain) and are also used in other multi-channel confocal papers (Walker et al., 1993; Grochmalicki et al., 1993). In related papers, (Bertero et al., 1990) further considers deconvolution approaches while (Defrise and De Mol, 1992) develops a Fourier-domain channel combination framework similar to the one used here.

Other confocally-based systems rely on multi-channel operation but to give different viewing orientations rather than improved photon efficiency. These techniques include SPIM (Huisken et al., 2004), MIAM (Swoger et al., 2003) and tomographic confocal microscopy (Heintzmann and Cremer, 2002). The cited papers give reconstruction methods for these systems. A disadvantage of these methods is that the data from each channel has to be aligned in the process of deconvolution. For a system using an array of detectors in the detector plane, this alignment is achieved automatically.

A detector array is used as a confocal detector in (Barth and Stelzer, 1994). If a detector array is used in place of a single pinhole, the photons discarded in a standard confocal instrument can be easily collected. Recent research has focused on developing such a detector array that is suitable for confocal microscopy (Pawley et al., 1996). In this section a multi-channel detector-array confocal system is defined, analyzed in terms of optical transfer functions and used to create simulated reconstructions at a variety of signal levels. This work differs from the publications referenced in that it gives simulated imaging results and examines the performance over a range of signal levels. This allows
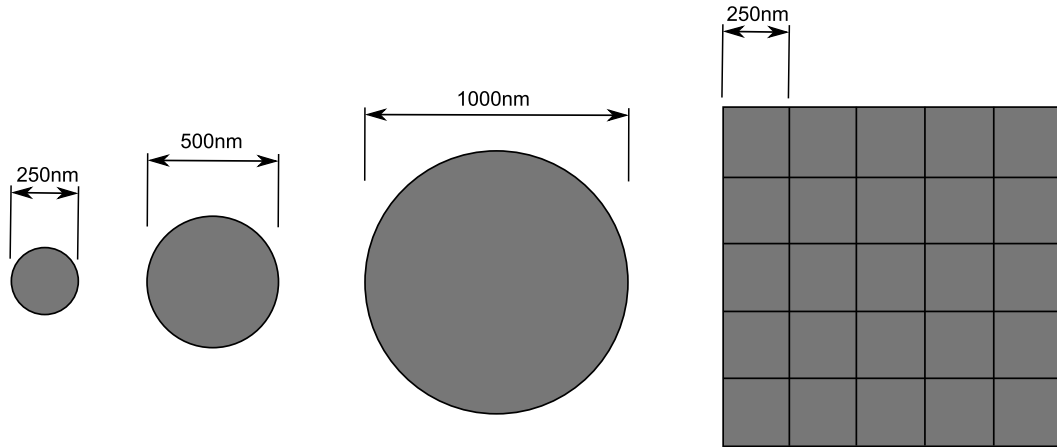
Figure 7·3: Scale diagram of the confocal detector shapes considered in this example. Three circular detectors of varying size are simulated along with a 5 × 5 grid of square detectors.

a quantitative analysis of the benefits of using a detector array as compared to a single pinhole.

The system considered uses lenses with a numerical aperture of 1.35 in a medium with refractive index 1.51. The excitation light is linearly polarized in the x-direction and has a wavelength of 488nm. Detection was carried out at 530nm and all spatial calculations were performed on a grid with spacings of 75nm axially and 22.5nm laterally. Four different detector apertures were considered and they are shown in Figure 7·3. Three standard circular pinholes are considered — they have demagnified diameters of 250nm, 500nm and 1000nm. The multi-channel instrument has a 5 × 5 array of square detectors with side-length of 250nm in demagnified units. This detector pattern is consistent with that proposed in (Pawley et al., 1996).

7.3.1   The Combined System's Optical Transfer Function

The optical transfer function for each detector is calculated according to the method presented in (Gu and Sheppard, 1992). The combined OTF for the multi-channel instrument was calculated according to Equation 7.4 and Equation 7.6. The 250nm-pinhole OTF was scaled so that the $\mathbf{k} = \mathbf{0}$ value was 1 and the other OTFs scaled so that they produced the
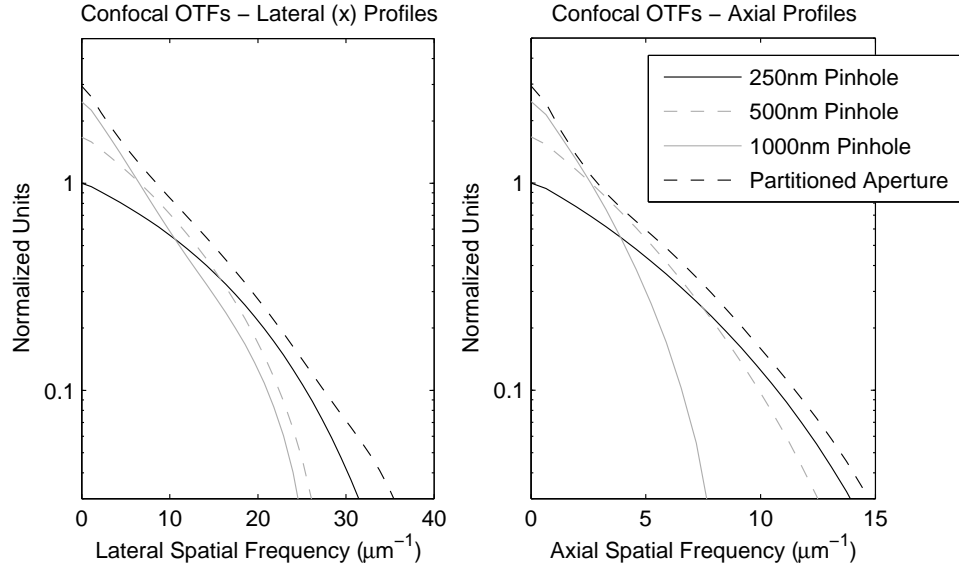
Figure 7·4: Lateral and axial profiles of the OTF magnitudes for a confocal system with a variety of pinhole sizes. The 'combined OTF' is also shown for a confocal system that uses an array of detectors in the detection focal plane.

same noise level. The magnitudes of these OTFs are displayed in Figure 7·4 and can also be regarded as SNRs according to Equation 5.25.

The standard-pinhole OTFs exhibit an intuitive behavior. At low spatial-frequencies the large pinholes have higher SNR values due to their superior photon efficiency. However, as the spatial-frequency is increased, the smaller pinholes perform better due to their ability to reject out-of-focus light. This illustrates the trade off present when selecting the pinhole size in confocal microscopy (Sheppard et al., 1991). It can be seen that the optimal pinhole size is dependent on the spatial frequency. In practice the pinhole size chosen will be one that performs well at the highest recoverable spatial frequency, which is itself determined by the noise level (Stelzer, 1998). The motivation for using a detector array is the idea that different pinhole sizes can be synthesized with post processing after the data is collected. This removes the need to pick a pinhole size before imaging but more importantly it allows the possibility of simultaneously synthesizing different pinhole sizes at different spatial frequencies. The 'Partitioned Aperture' plot shows that the detector array can indeed
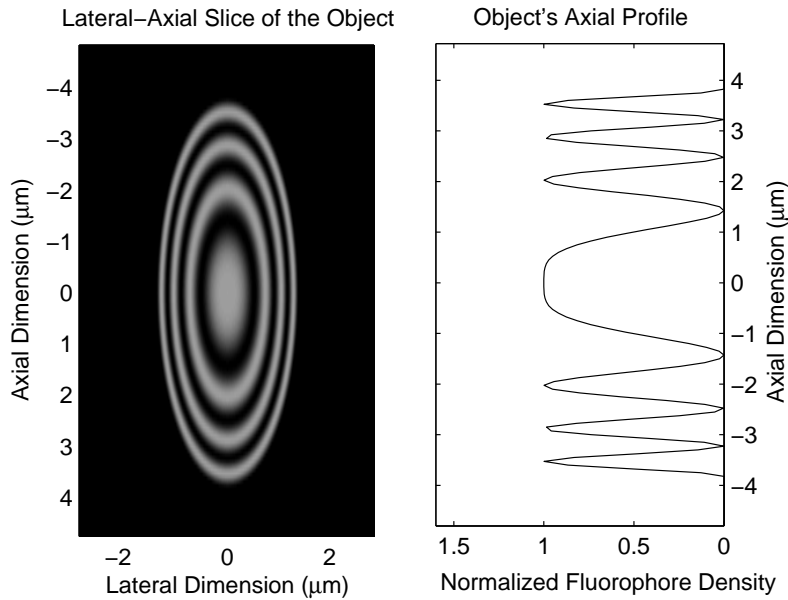
Figure 7·5: An image of an x-z plane in the confocal test object and a plot of the object's axial profile. The object is cylindrically symmetric around the z axis.

produce this kind of behavior. The combined OTF for this system outperforms all of the pinholes at all spatial frequencies.

7.3.2   Simulated Reconstructions

As in previous sections, a test object will be defined and used in simulated imaging to test the performance of the systems examined. The test object used is shown in Figure 7·5. This is similar to the test objects used in the previous chapter but has been elongated in the axial direction. This is so that the outer layer spacings test the limits of the instruments' resolution in both the axial and lateral directions.

In order to perform reconstructions of the object from data it is necessary to define the 'Single-Channel Inversion' block seen in Figure 7·1 (the rest of the system has already been defined). The Wiener filter (Wiener, 1964) will be used for this purpose. This standard inversion filter is shown below and is defined by a MMSE rule under a stationary noise

assumption.

$$\hat{O}(\mathbf{k}) = \frac{H^*(\mathbf{k})}{|H(\mathbf{k})|^2 + \alpha^2} D(\mathbf{k}) \qquad (7.7)$$

Ideally the α term is related to the signal to noise ratio and is dependent on $\mathbf{k}$ and the object. This is not practical as the object will typically be unknown. For this reason α is simplified to a constant and regarded as a regularization parameter. Any resulting negative values in $\hat{o}(\mathbf{r})$ will be set to zero as the fluorophore density is known to be non-negative.

Data was generated from the object for each instrument and scaled to a variety of levels before being realized as a Poisson random variable. For each data set a reconstruction was performed for a wide range of α values. The α that minimized the square difference between the reconstruction and the true object was then chosen as the optimal value of the regularization parameter for the corresponding instrument and signal level. In practice the regularization parameter cannot be selected in this manner since the object is unknown but for the purposes of comparisons across instruments the optimal α value will be used. Two illustrative reconstructions are shown in Figure 7·6. They are reconstructions from the multi-channel instrument and the best (in the sense that it gives the smallest error) standard pinhole for the same signal level. Throughout this chapter instruments at the same signal level will be compared and by this it is meant that the same number of photons fall on the detector plane.

The difference in the reconstructions in Figure 7·6 is small but clear. The multi-channel instrument quantitatively matches the fluorophore density better as can be clearly seen in the axial profiles. The peaks and valleys are also more clearly resolved and the contours in the lateral-axial slice are smoother. This indicates a superior imaging by the multi-channel instrument.

The imaging quality can also be measured quantitatively. For each signal level and each instrument 100 realizations of the Poisson noise were calculated. The object was then estimated from these data sets using the optimal α value. The average square errors are plotted in Figure 7·7. Again, the expected behavior is observed. The larger pinholes perform better at lower signal levels since it is important to maximize the SNR by col-
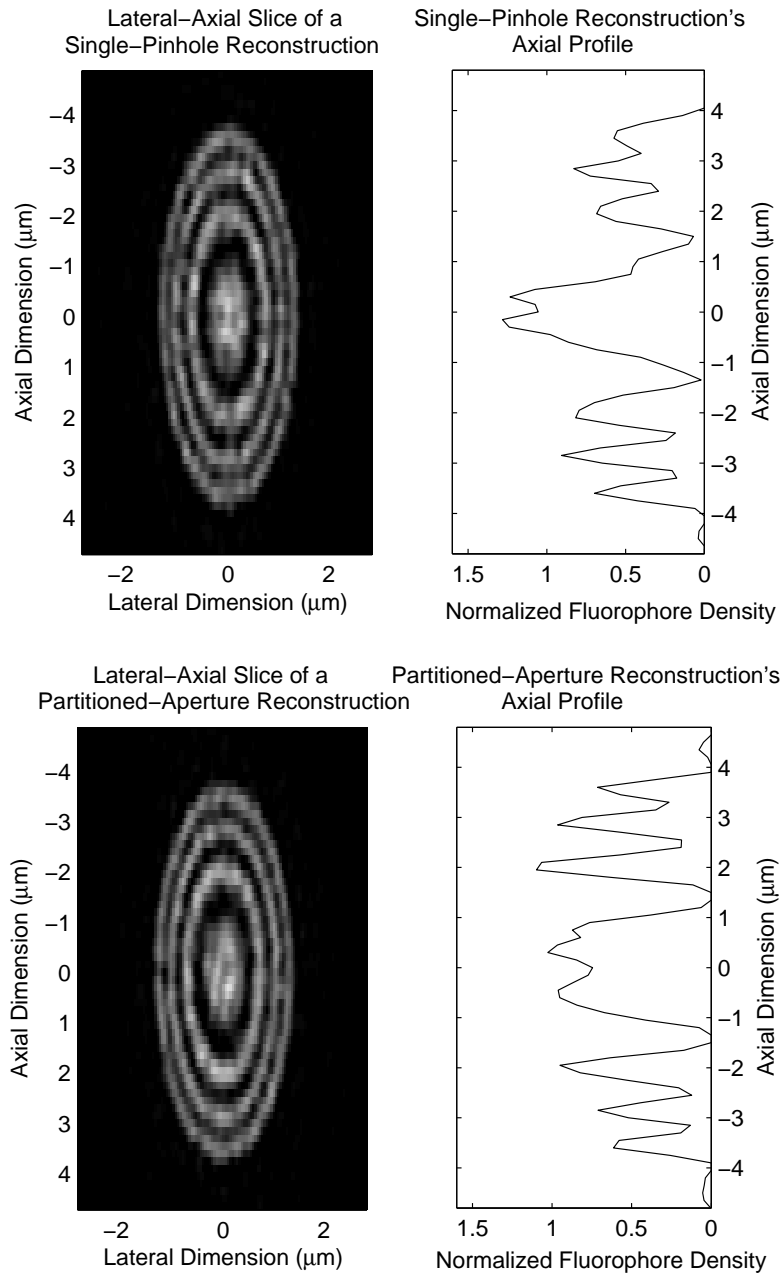
Figure 7·6: Reconstructions for a confocal system with a 250nm pinhole and for a confocal system with an array of detectors in the detection focal plane. An x-z image and an axial profile are shown. The same signal strength was used in each case and corresponds to a peak expected count of 20 in the single-pinhole data.

Confocal Reconstruction Error at a Range of Signal Levels
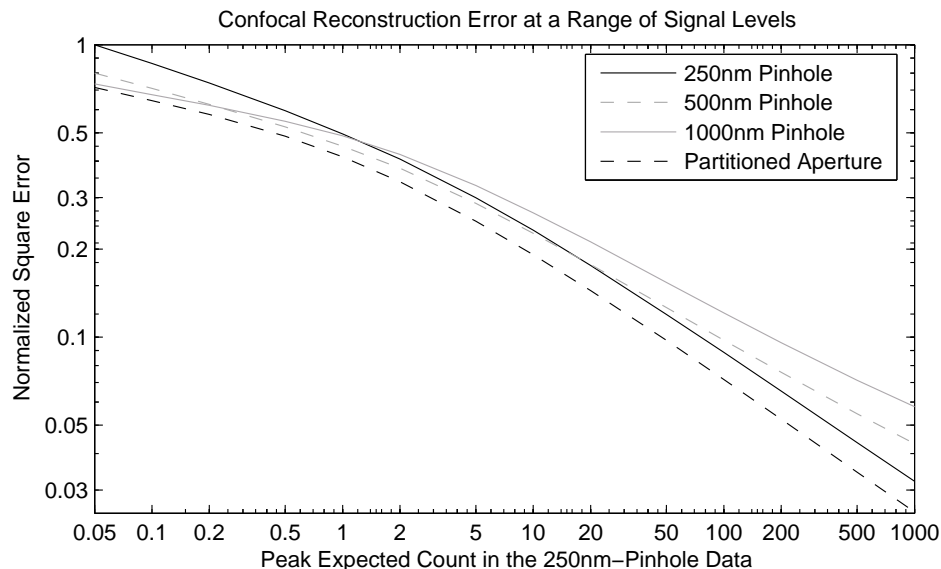
Figure 7·7: The confocal reconstruction square-error as a function of signal level. The average reconstruction error from 100 trials was calculated for each of the signal strengths displayed on the x axis. Each trial had a different Poisson noise realization and the standard deviation of the reconstruction error was less that 1.3% of the mean in all cases.

lecting as many photons as possible. Smaller pinholes perform better as the signal level increases and the SNR at lower spatial frequencies is high. In this case it is important to raise the OTF values at higher spatial frequencies by using a narrower pinhole (i.e. rejecting out-of-focus light). Again, the multi-detector instrument with a partitioned detection aperture performs better than all the other systems at all signal levels. Quantitatively, the error suggests that the multi-channel system performs as well as a single-channel system collecting up to as much as 50% more photons.

The combination filter $\eta^{(m)}(\mathbf{k})$ gives a weighting to the mth detector element at spatial frequency $\mathbf{k}$. At a given spatial frequency the weighting can be plotted as a function of detector position to give an artificial aperture. Some examples of these are shown in Figure 7·8. Here the weighting patterns are shown for three spatial frequencies on the axial axis. At zero spatial frequency all of the detectors are equally weighted so that all of the collected photons are simply averaged. Since the $\mathbf{k} = \mathbf{0}$ component of the object is simply the total amount of light emitted, this detector weighting makes sense. As the axial spatial
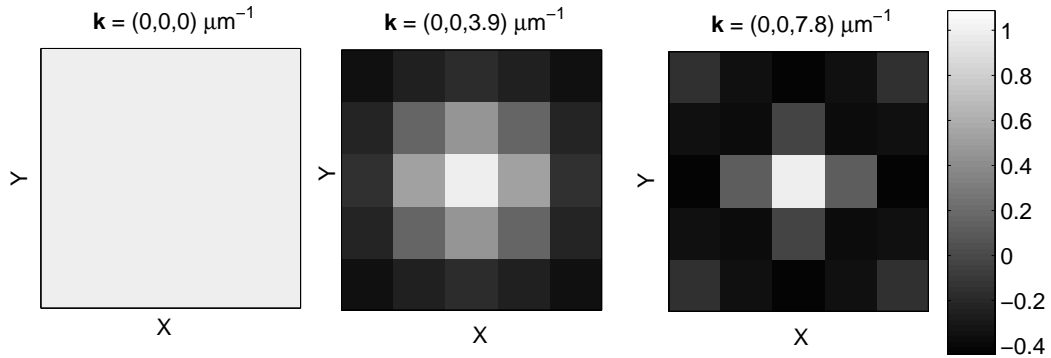
Figure 7·8: The detector weightings for several spatial frequencies in the partitioned-detector instrument. Note the detector weights have been normalized to have a maximum of 1 in each image and the $\mathbf{k} = (0,0,0)$ case on the left has a uniform weighting.

frequency is increased an increasingly confined pinhole is observed. Again this is intuitive as detecting a higher spatial frequency requires a narrower pinhole while the aperture can be opened at lower spatial frequencies to improve the SNR.

The above simulations have shown how a multi-element detector can be used to improve the imaging quality in confocal microscopy. However, it should be noted that this moderate gain was demonstrated assuming photon-counting limited noise and an equal quantum efficiency in all detectors. In practice the detector array can be expected to have different noise and efficiency properties than a single-channel detector (typically a Avalanche Photodiode (APD) or Photomultiplier Tube (PMT) in confocal microscopy). Before attempting to implement such a multi-channel confocal microscope it is therefore necessary to carefully analyze the expected performance for the proposed detector, as in (Sheppard et al., 1992). The results presented here are simply what can be expected if the detector characteristics are equal. Proposed confocal detector arrays have been shown to have good noise characteristics in comparison to standard detectors (Pawley et al., 1996; Coates et al., 2003).

7.3.3 Optimal Confocal Apertures

The weighting patterns shown in Figure 7·8 show how the multi-channel instrument essentially implements an optimized pinhole shape at each spatial frequency. Because of this it performs better than a fixed physical aperture which must apply the same pinhole shape to all spatial frequencies. Additionally, the weightings shown in Figure 7·8 are all real but for off-axis spatial frequencies complex values of $\eta^{(m)}(\mathbf{k})$ were given by Equation 7.6. A complex physical aperture cannot easily be implemented as the light falling on the detector is incoherently summed (see Equation 4.12). In this sense the multi-channel confocal microscope is more flexible than a single channel instrument or even the proposed two-channel systems using optimized complimentary apertures (Walker et al., 1993; Verveer and Jovin, 1998).

More general optimal apertures can be calculated by using a very fine discretization of the detector plane (Davis et al., 2005c). The same system as before is considered but now rather than having detection elements of 250nm × 250nm (in demagnified units), the new system will have detection elements of 10nm × 10nm. Some resulting detection patterns are shown in Figure 7·9. The spatial frequencies selected in this figure show both real ($\mathbf{k}$ on the axial axis) and complex apertures. As with the earlier results it can be seen that the aperture tends to get smaller for higher spatial frequencies. The aperture for $\mathbf{k} = (18, 52, 2.6)\mu\mathrm{m}^{-1}$ shows phase and directional structure introduced by the fact that lateral frequencies are being imaged. The orientation of these phase and magnitude structures are consistent with the lateral frequency imaged.

Looking at the plots of Figure 7·9, it can be seen that the resulting apertures vary relatively slowly compared to the spacing of the detectors. This would indicate that the 10nm detector size provides a reasonable approximation of the ideal case — an infinitely finely partitioned detector plane. Thus the 10nm detector array provides enough flexibility to approximate the optimal collection of the emitted photons. However, at this detector size there would be difficulties in fabricating the detector array. An objective with 100× magnification would mean that the individual detectors would have to be a challenging
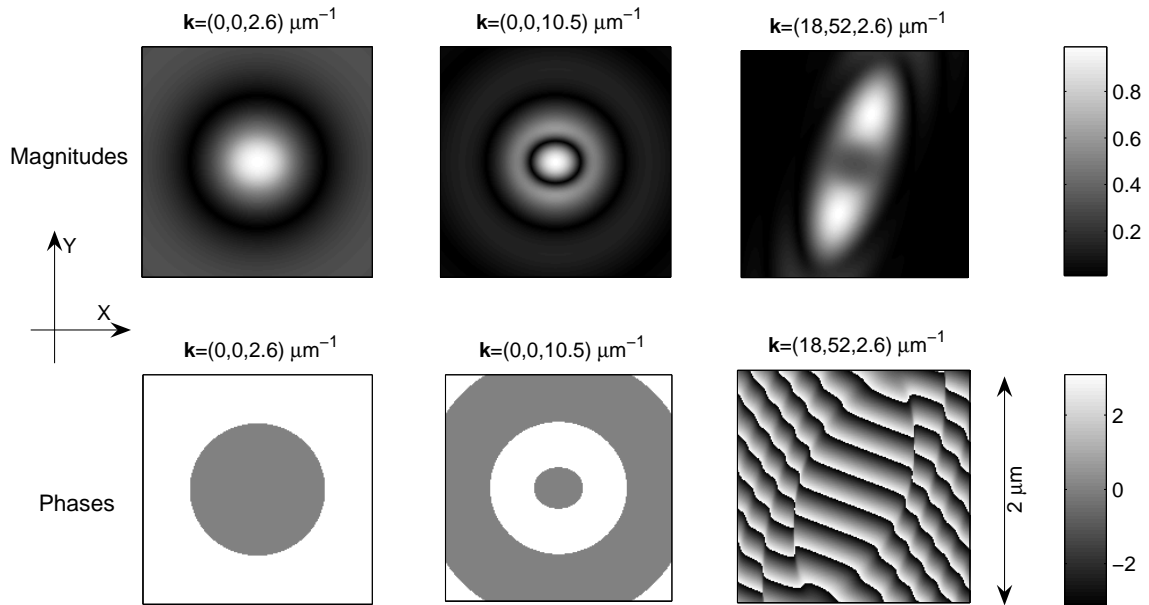
Figure 7·9: Detector weightings for several spatial frequencies in a finely-partitioned detector instrument. The detector weights have been normalized to have a maximum magnitude of 1 in each image. These plots can be regarded as optimal complex detection apertures. Magnitudes are displayed on the top row and phases on the bottom. The detector dimensions are 2μm by 2μm.

1μm in size. Supposing such an array could be constructed, it would be reasonable to ask what the payoff would be in terms of imaging.

In Figure 7·10 the combined OTF for the 10nm-detector system is compared to that of the 250nm-detector system examined earlier. The finely-partitioned system does perform better but not by a large amount. At very low spatial frequencies the performance is higher due to a larger detector area and at higher frequencies the smaller individual detector size allows a better SNR. To get a factor of two improvement in the SNR, one has to consider a $\mathbf{k}$ where $H_{\mathrm{Comb.}}(\mathbf{k})$ is less than 40 times smaller than $H_{\mathrm{Comb.}}(\mathbf{0})$ for the optimal system. This implies that the finely-discretized system has very little advantage over the 5×5 array unless the signal strength is very high.
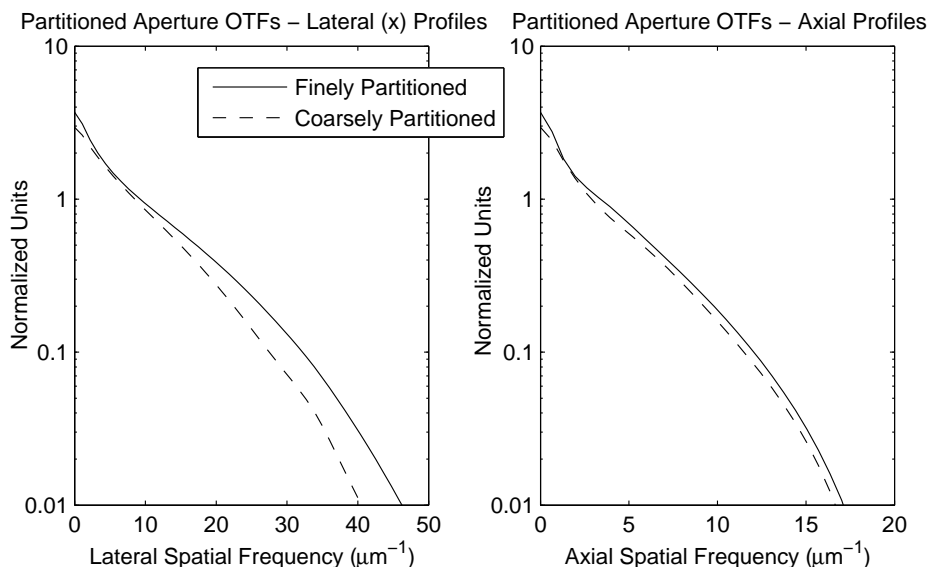
Figure 7·10: Lateral and axial 'combined OTF' magnitude profiles for partitioned detector confocal systems. Coarsely (250nm elements) and finely (10nm elements) partitioned cases are considered.

7.4   Two-Channel 4Pi Microscopy

Multi-channel techniques can also be applied to 4Pi systems. Looking at Figure 4·1 it can be seen that only one of the two output ports of the beamsplitter has detection apparatus attached to it. This means that emitted light is discarded in the second port. While the amount of light wasted in this manner can be minimized by ensuring that the detection phase ϕ_d is zero, there is still a significant amount of light lost. If this light is collected it can be regarded as a second channel and used to improve the image quality by the methods described in this chapter. This idea will be tested through numerical example modeled on the instrument used at the Deutsches Krebsforschungzentrum (DKFZ) — this microscope is described in (Gugel et al., 2004).

7.4.1   The Combined System's Optical Transfer Function

A two-photon-excitation 4Pi system (Hell and Stelzer, 1992b; Hell et al., 1997) will be considered here. The multi-photon excitation is used to reduce sidelobes in the PSF which in turn makes the valleys within the OTF support (see Figure 4·6) less pronounced and

allows for better imaging (Gu and Sheppard, 1995; Nagorni and Hell, 2001b; Gugel et al., 2004). The signal level for biological samples in the two-photon 4Pi system at the DKFZ tended to be very low (on the order of one to a few counts per pixel). This strongly motivates collecting the second-channel light. Additionally, the scanning method used at the DKFZ meant that the focal point for the second channel moved on the detector plane and thus was collected using a large detector. The net effect is the same as using an open detector $s(\mathbf{r}) = 1$ — for this reason the second channel is also referred to as a non-confocal channel. A non-confocal, two-photon 4Pi instrument is presented in (Gugel et al., 2004).

The two-photon 4Pi system considered used a x-polarized excitation light at 870nm and detected at 515nm. The lenses used had a numerical aperture of 1.39 in a medium with an index of refraction of 1.51. The pinhole for the standard channel is circular with a demagnified diameter of 240nm while the non-confocal channel had a pinhole diameter of 4.8μm. The excitation phase was zero ($\phi_e = 0$) and the detection phase in the standard channel was also zero. This means that the detection phase in the non-confocal channel was 180°. The calculations were carried out on a spatial grid with spacings of 10nm axially and 30nm laterally.

Profiles of the resulting OTF magnitudes of the two channels and the combined OTF are shown in Figure 7·11. Again, these are displayed so that the noise is equal in each system and the OTFs can therefore be regarded as SNRs. While the second channel collects a significant number of photons (as can be seen by the high value of the OTF at $\mathbf{k} = \mathbf{0}$) this data is not of high quality since the OTF drops off rapidly. This means that the combined OTF is only raised marginally over the standard-channel OTF. The effect of this on imaging can be tested through simulations.

7.4.2 Richardson-Lucy Simulation Results

As mentioned earlier, there are many possible algorithms that can be applied to reconstruct the object from multi-channel data. In this section the Poisson-based Richardson-Lucy algorithm will be used. This means that the data sets are not explicitly combined into a
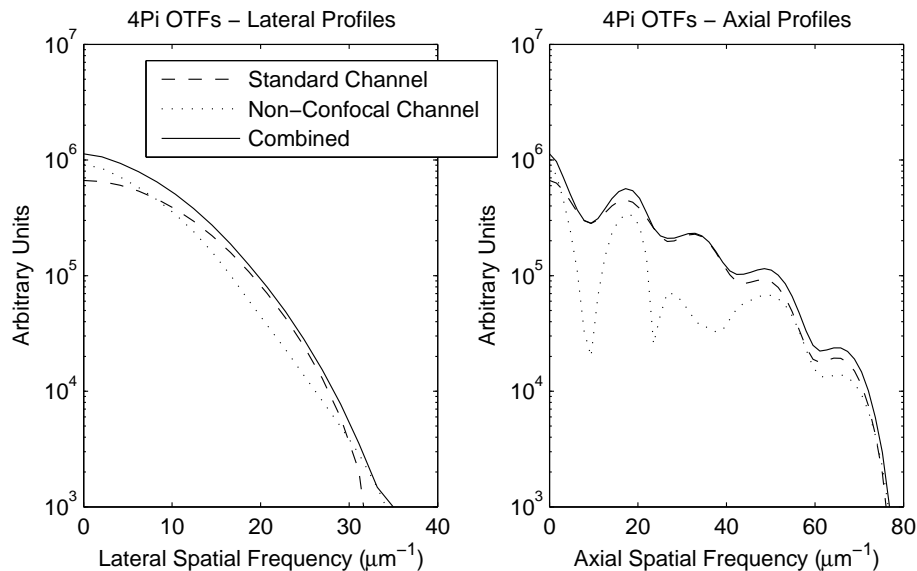
Figure 7·11: Lateral and axial profiles of the normalized OTF magnitudes for a two-photon 4Pi Type C system — each of the channels and the combined system are shown.

single data set before processing. The results of these reconstructions will give not only a characterization of the multi-channel 4Pi system but also a point of comparison between two different multi-channel reconstruction approaches.

The Richardson-Lucy algorithm is expressed for matrix equations in Equation 3.18. However, the algorithm only requires knowledge of the forward and adjoint operators of the system. For a linear shift-invariant system the forward operator is simply convolution with the PSF $h(\mathbf{r})$ while the adjoint operator is simply convolution with $h^*(-\mathbf{r})$. So a two-channel system can be defined by the following forward model.

$$d^{(1)}(\mathbf{r}) = h^{(1)}(\mathbf{r}) * o(\mathbf{r}), \qquad d^{(2)}(\mathbf{r}) = h^{(2)}(\mathbf{r}) * o(\mathbf{r}) \qquad (7.8)$$

The definition of the Richardson-Lucy algorithm is easily applied to give the following definition of an iteration of the algorithm. This iteration takes an estimate of the object
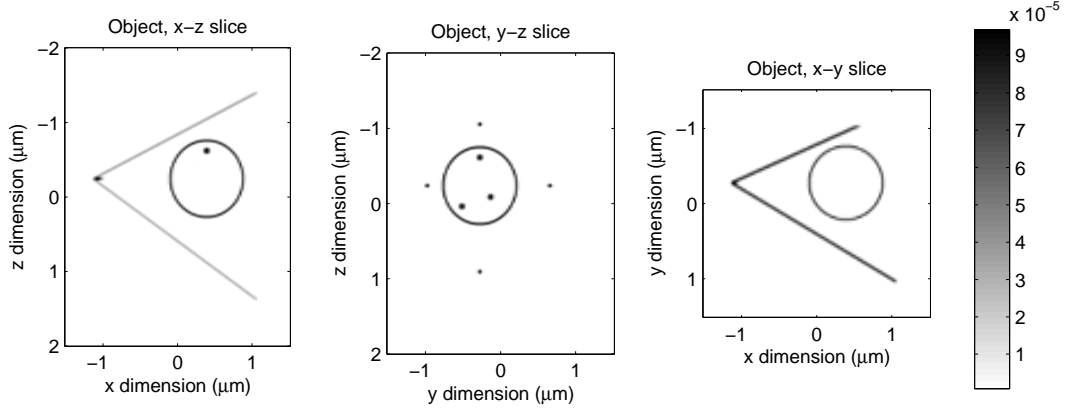
Figure 7·12: Images of differently oriented planes within the test object used in the 4Pi system simulations.

$\hat{o}^{(v)}(\mathbf{r})$ and produces the next estimate in the algorithm $\hat{o}^{(v+1)}(\mathbf{r})$.

$$\hat{o}^{(v+1)}(\mathbf{r}) = \hat{o}^{(v)}(\mathbf{r}) \left\{ \left[h^{(1)}(-\mathbf{r}) \right]^* * \frac{d^{(1)}(\mathbf{r})}{h^{(1)}(\mathbf{r}) * o^{(v)}(\mathbf{r})} + \left[h^{(2)}(-\mathbf{r}) \right]^* * \frac{d^{(2)}(\mathbf{r})}{h^{(2)}(\mathbf{r}) * o^{(v)}(\mathbf{r})} \right\} \tag{7.9}$$

Here it has been assumed that the PSFs satisfy the following relationship.

$$\int \left[h^{(1)}(\mathbf{r}) + h^{(2)}(\mathbf{r}) \, d\mathbf{r} \right] = 1 \tag{7.10}$$

An arbitrary system can easily be scaled so that this condition is met.

The test object will be the same as that used in (Nagorni and Hell, 2001a). This object consists of a sphere with several point objects inside it and four line objects outside. This object is a model of a cell with features inside and filamentary structures surrounding it. Some cross sections of this object are shown in Figure 7·12. A three-dimensional rendering of the object is shown in Figure 7·13.

Simulated data with Poisson noise can be generated from this object and Richardson-Lucy reconstructions found from that. In these test simulations the object is known so the error in the reconstruction can be calculated. In all cases the algorithm was stopped at the iteration that produced the minimum error. This stopping criterion is not practical
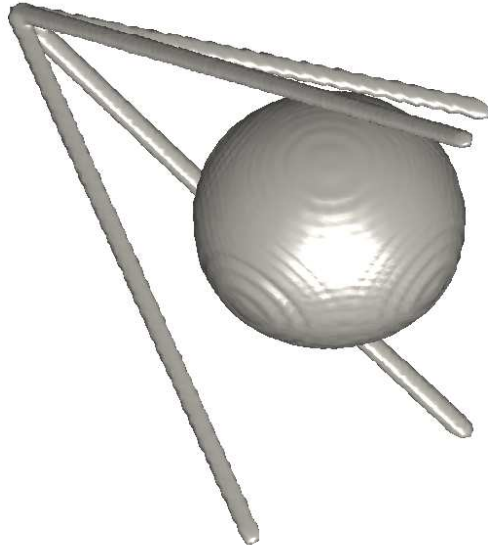
Figure 7·13: A three-dimensional rendering of the test object used in the 4Pi system simulations. Note that the features inside the sphere are not visible.

in real systems where the object is unknown but it does provide a good way to compare systems in this investigation. The error measure used was the I-divergence (Csiszár, 1991) which is matched to Poisson statistics and the Richardson-Lucy algorithm. The previous reconstructions in this chapter were based on a MMSE estimation approach and thus the square error was the appropriate measure in that case.

Some reconstructions from the 4Pi system are shown in Figure 7·14. The maximum expected count in the single-channel data was only 0.5 — this shows that good images can be generated even from very noisy data. This data level was displayed as it gave some of the clearest differences between the one- and two-channel reconstructions. The improvement is quite small but, as shown in Figure 7·15, visible. In this case one of the filamentary structures isn't resolved from the cell in the single-channel reconstruction, while it is in the two-channel reconstruction.
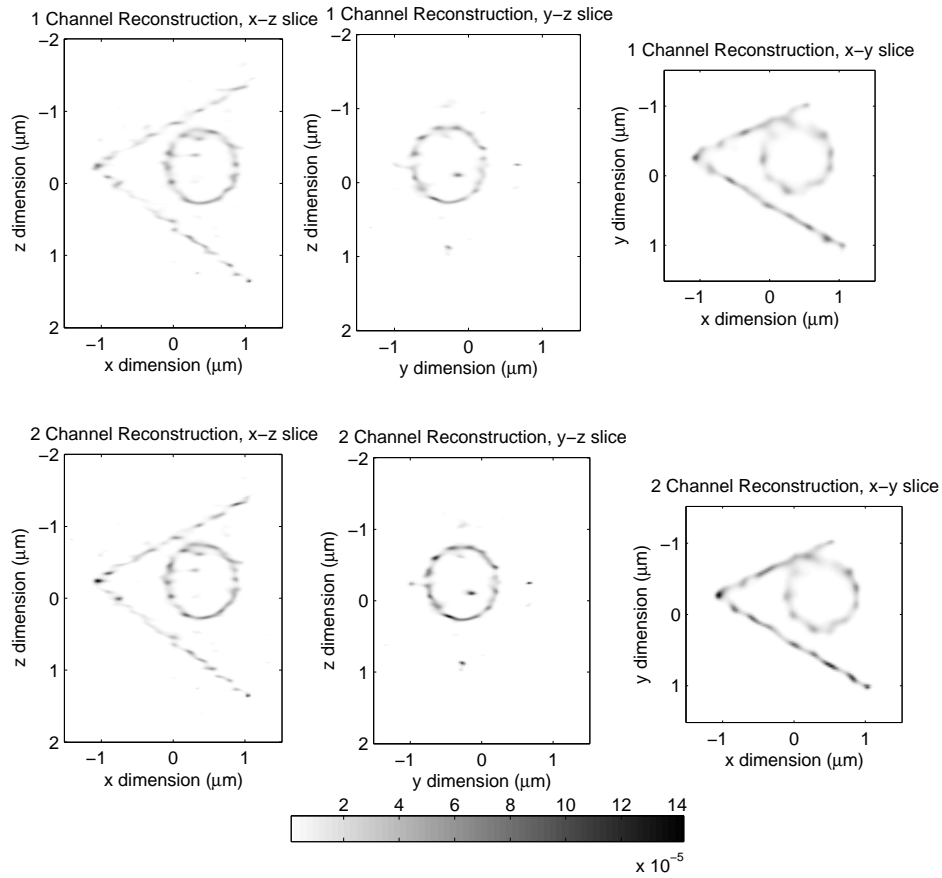
Figure 7·14: Reconstructions of the test object from single-channel and two-channel 4Pi data. The data has a peak expected count of 0.5 in the standard channel. The planes shown are the same as those in Figure 7·12.
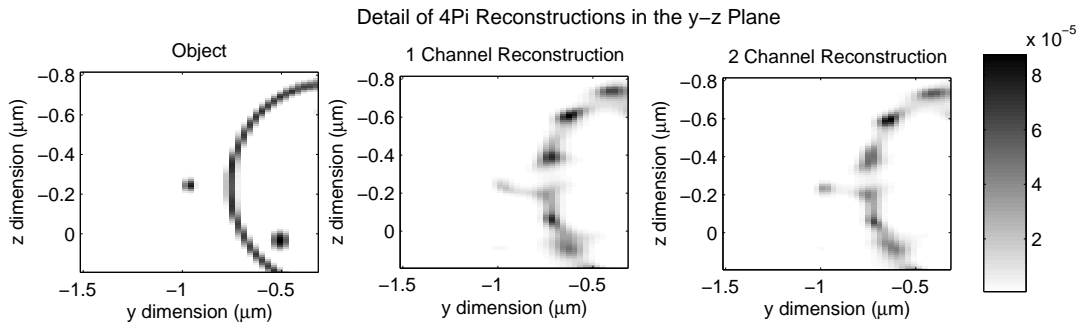
Figure 7·15: Detail for the 4Pi object, single-channel reconstruction and double-channel reconstruction.
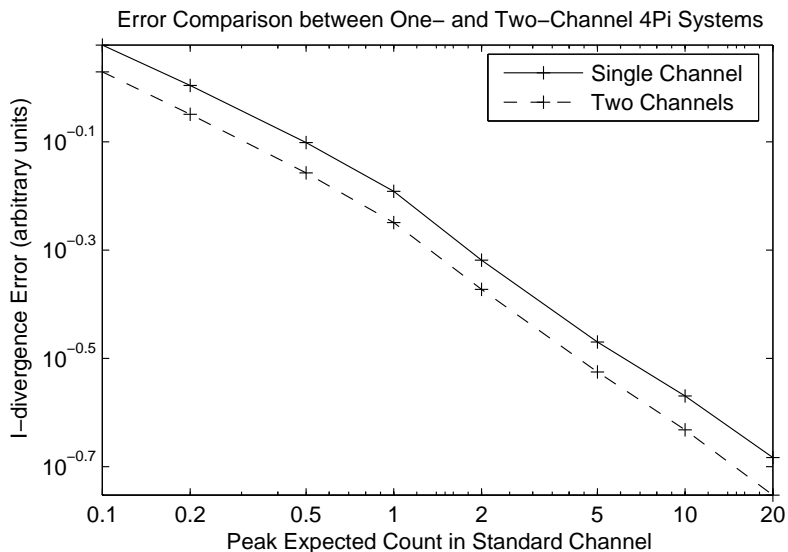
Figure 7·16: The 4Pi reconstruction I-divergence error as a function of signal level.

As in the multi-channel confocal case, the error can also be plotted as a function of signal level. This is done in Figure 7·16. Due to the length of time taken to perform the Richardson-Lucy iterations, each data point is from only a single reconstruction (c.f. Figure 7·7 where each point was the average of 100 reconstruction errors). However, the smooth plots indicate that the calculated errors are probably not highly variable. The error plots indicate that the extra channel has been usefully incorporated into the reconstruction as the error level drops at all signal levels. Like the multi-channel confocal example, the two-channel 4Pi instrument achieves reconstruction errors similar to a single-channel 4Pi instrument collecting approximately 50% more photons. The roughly constant gap between the two curves in the logarithmic plot of Figure 7·16 corresponds to an error reduction of about 12% when the second channel is included in the reconstruction.

7.5   Summary

In this section it has been shown that multi-channel microscopes can be used to beneficially collect light that, in single-channel systems, is considered to be of low quality and is typically discarded to improve imaging quality. Collecting this low quality light in a

separate channel allows it to be included in the reconstruction in a useful manner — this is significantly different to a single-channel instrument where physically altering the system to collect the additional light also determines how it affects the reconstruction (which is generally in a detrimental manner). A simple Fourier-domain method is developed to perform the multi-channel reconstruction and a measure of the resulting SNR is naturally defined by this method through the concept of a 'combined OTF'. As shown through a Richardson-Lucy example, other image reconstruction methods can be applied to multi-channel data with similar results.

The multi-channel analysis is applied to a SSFM system. The example shows that the addition of the SSFM phase offsets produces only a very minor increase in SNR at some of the higher spatial frequencies. This is due to the fact that certain channels collect less photons (which gives less noise) but still pass the high axial spatial frequencies well. However, for the great majority of spatial frequencies, a system without phase offsets and collecting across the whole spectrum performs better than the SSFM system. It should be noted that maintaining known phase offsets across a broad detection spectrum is a difficult task and that collecting at only a single wavelength produces an SNR profile only a little lower than a system using spectral detection. The analysis used also assumed exact knowledge of the fluorophores' spectral emission envelope. These simulations essentially indicate that the difficult engineering challenges involved in making a general spectral 4Pi or SSFM system can be expected to produce only a small gain in performance.

The multi-channel approach was also applied to a confocal instrument using a detector array instead of a pinhole and single detector. This has been proposed in the literature before but simulations investigating the imaging performance changes have not been published. The results show how a multi-detector system outperforms standard confocal systems at all signal levels and all pinhole sizes (assuming otherwise equal detection characteristics). The multi-channel processing algorithm is also interpreted in terms of synthesizing different aperture shapes at different spatial frequencies. The results provide an intuitive framework for analyzing the multi-detector confocal instrument but while the

simulation results do show improvement, the gain is small. In light of the numerous practical difficulties involved in constructing, calibrating and operating a confocal microscope (Pawley, 2000), the multi-detector improvement would only be significant in a very well designed and operated system. It does however provide a method for incorporating the recent developments in the field of high-efficiency, low-noise detector arrays.

A two-channel 4Pi system is also considered where the second channel is found by detecting the light coming out of the second channel of the beamsplitter. The combined OTF also shows a limited improvement in the realized SNR. Simulated imaging is then shown but using a Richardson-Lucy deconvolution method rather than the Fourier-domain method developed here. The results are similar in that a small gain in imaging performance is seen and a quantitative analysis also shows a consistent improvement in imaging quality at all signal levels. In this example it could also be claimed that the additional complexity involved in constructing and calibrating the second channel may outweigh the performance gain allowed.

Chapter 8

Summary and Future Work

To conclude this work a very brief chapter-by-chapter summary of the most important results will be given, followed by suggestions for future work relating to the contents of this dissertation.

8.1   Chapter-by-Chapter Summary

Chapter 1: Introduction

This chapter very briefly motivates fluorescence microscopy and describes several techniques and their imaging capabilities. More time is devoted to widefield, confocal and 4Pi imaging as these techniques are directly related to the work presented in this dissertation. SSFM is described along with its early promising results.

Chapter 2: SSFM Modeling Framework

This short preliminary chapter outlines some of the underlying principles used in SSFM modeling and simulation. Specifically, the relationship between discrete and continuous models is described and noise statistics are specified. A stochastic justification for the measurement of interferometric data from the incoherent fluorescent source is also presented.

Chapter 3: SSFM Image Reconstruction and Performance Evaluation

Numerical simulations are used to provide examples of SSFM image reconstruction in this chapter. Several approaches are applied to an axial test object and an inversion method is developed for objects that are known to consist of only a few object components. It is seen

that the prior knowledge inherent in the later case allows a much greater reconstruction accuracy. The examples also indicate that the high accuracy of the early SSFM results is also dependent on this prior knowledge. Comparing SSFM reconstructions to similar examples for confocal, I^5M and 4Pi systems indicates that SSFM and 4Pi microscopy have similar imaging capabilities.

Chapter 4: Relation of SSFM to Known Imaging Systems

This chapter further explores the connections between SSFM and known microscopy systems. It is shown that scanning SSFM is essentially 4Pi system but with spectral detection, variable interference properties and a possible mismatch between objective foci. To characterize the instruments mathematically a linear, shift-invariant systems analysis is used. It is shown how the PSF of many microscopy systems, including SSFM and 4Pi, can be calculated from the diffraction pattern of a lens. An inspection of the resulting OTFs also indicates that 4Pi and SSFM systems have a similar achievable resolution.

Chapter 5: Analysis of Optical Transfer Function Support

While the previous chapters have relied on illustrative examples, this chapter includes a rigorous reasoning as to why SSFM cannot have an OTF support beyond that of a similar 4Pi system. The OTF support determines the maximum spatial frequencies present in the data and thus the maximum achievable resolution (without prior knowledge of the object). The OTF-support result relies on a Fourier-domain construction of the OTF based on the Fourier transform of the lens diffraction pattern. A similar analysis can also be applied to numerous other microscopy systems including widefield, confocal and 4Pi microscopes. It is seen that such systems have a strictly limited OTF support and approaches such as pupil-filter apodization and SSFM cannot extend it. However, there are potential gains in terms of the signal-to-noise ratio.

Chapter 6: Relaxing Scanning Requirements with Multi-Channel Microscopy

The fact that SSFM systems can perform axial imaging of thin objects without scanning suggests that scanning requirements can be reduced using multi-channel systems such as SSFM. This chapter provides a mathematical framework for doing exactly that and simulations are performed to demonstrate the concept. A SSFM example and a multi-detector confocal example are given and shown to allow sub-Nyquist-rate scanning. The motivation for this approach is the hope that imaging time may be reduced if fewer sample points can be collected. However, an analysis of performance in the presence of noise shows that multi-channel sub-Nyquist scanning can do no better than multi-channel Nyquist-rate scanning but with the dwell time reduced to give the same imaging time.

Chapter 7: Suppressing Noise with Multi-Channel Microscopy

This chapter considers the performance of multi-channel microscopy systems operating at the Nyquist rate. The signal processing methodology developed combines the multiple data sets in a noise optimal manner and also defines a measure of the SNR as a function of spatial frequency. A SSFM example is considered and a very small SNR gain (compared to a comparable 4Pi system) can be seen at high spatial frequencies. Lower spatial frequencies' SNR tends to suffer. A multi-detector confocal example shows a small but clear performance enhancement over single-channel instruments and similarly a two-channel 4Pi system performs better than a single-channel 4Pi microscope.

8.2   Future Work

While this work has presented a thorough analysis of SSFM, developed its relation to other microscopy systems and provided a clearer understanding of its capabilities and limitations, there is still significant further work that can be pursued. It has been demonstrated that SSFM can provide axial imaging of a limited-extent object without employing any axial scanning. This provides a significant instrument simplification over scanning techniques. As discussed in Section 6.4 there is also the possibility that a multiplex-advantage

effect will provide a SNR benefit as signals from multiple spatial positions are collected simultaneously. For these reasons a future investigation into the quality of SSFM imaging for sparse and/or limited-axial-extent objects is proposed. It should also be noted that evanescently-excited SSFM provides a suitably axially-limited fluorescent object due to the narrow excitation profile.

Another area of future research includes physical demonstrations of the theoretical results presented in this dissertation. This would include developing working prototypes of, for example, the multi-detector confocal microscope and the evanescently-excited SSFM system. These first-generation instruments would provide a proof of concept and also allow biological imaging to be carried out. The search for prospective imaging applications for SSFM and related technologies is also an ongoing area of investigation.

Appendix

Notation and a List of Variables

The mathematical notation throughout this dissertation has been kept consistent. Except for a few general-purpose indices, no variable, typeface, accent, etc. has been used to denote more than one term or concept. This section gives a glossary of the notation used.

General Notation

The list below outlines the use of typefaces, accents, brackets, subscripts, superscripts, etc.

- A boldface variable is a vector in the sense that it is a collection of numbers or variables, not in the sense that it satisfies a set of mathematical axioms. For example, $\mathbf{r}$ is the position vector and consists of three numbers giving x, y and z locations.

- Three-dimensional vector functions are often used and will be denoted with a small arrow over the function. For example, $\vec{a}(\mathbf{r})$ is a vector function that varies with the vector $\mathbf{r}$. So $\vec{a}(\mathbf{r})$ is made from three scalar functions of $\mathbf{r}$, one for each component of the vector output.

- A prime $'$ is added to a continuous variable to produce a dummy variable in an integral. For example, in Equation 2.5 t' is used to integrate over time.

- The accent $\hat{\ }$ is used to denote an estimate. For example, $\hat{\mathbf{o}}$ means estimate of $\mathbf{o}$.

- An overbar $\bar{\ }$ is used to indicate sampled data when sampling effects are being explicitly considered.

- Subscripts are used to differentiate between two different uses of the same variable.

For example, λ indicates wavelength while λ_e and λ_d denote excitation and detection wavelengths respectively.

- A subscript directly outside a bracketed expression has a different meaning. In this case the subscripted term indicates an individual element in a vector or matrix. For example, $(\mathbf{d}_e)_u$ means u^{th} component of $\mathbf{d}_e$ and $(C)_{uv}$ is the element found at the intersection of the uth row and the vth column of the matrix C. Additionally, this notation is used to identify x, y and z components in functions that depend on the 3×1 vector $\mathbf{r}$. For example, $(\mathbf{k})_z$ is the z component of $\mathbf{k}$.

- The brackets (), [] and {} do not have different meanings. All are used to group terms and will typically be used together in expressions where multiple brackets are required.

- A bracketed superscript indicates the index of a series of terms. For example, $h_e^{(j)}(\mathbf{r})$ is the jth excitation point spread function.

- Terms appearing after a semicolon ; in a function argument can be considered as parameters while those before the semicolon can be considered as variables. For example, $d(\mathbf{r}; \lambda)$ represents the data as a function of the position $\mathbf{r}$ collected at the fixed wavelength λ.

- Fourier transforms can be taken either spatially or temporally. A temporal function will have a dot $\cdot$ above it while its temporal-frequency counterpart won't. A spatial function will be lowercase while its spatial-frequency counterpart will be capitalized. For example, $\dot{h}(t, \mathbf{r})$ is taken to $H(\nu, \mathbf{k})$ by a three-dimensional spatial Fourier transform and a temporal Fourier transform. Note that the temporal frequency variable ν is calculated using the Fourier kernel $e^{-i2\pi\nu t}$ while the spatial frequency variable $\mathbf{k}$ is calculated using the kernel $e^{-i\mathbf{k} \cdot \mathbf{r}}$.

List of Variables

A list of variables is presented below. This list is presented in alphabetical order with Greek and miscellaneous variables following. Common functions, constants and notation (such as $\sin(\cdot)$, π, e and the use of $^{-1}$ to indicate inverse function) are not listed unless they are different from standard definitions. Functions are listed with their most common arguments although it should be noted that these are not necessarily the only possible arguments. This list is exhaustive up to variables distinguished by subscripts (e.g. $h(\mathbf{r})$ is listed as a point spread function but the point spread functions for widefield ($h_{\text{Wide.}}(\mathbf{r})$), confocal ($h_{\text{Conf.}}(\mathbf{r})$), 4Pi ($h_{\text{4Pi}}(\mathbf{r})$) and SSFM ($h_{\text{SSFM}}(\mathbf{r})$) systems are not listed individually). In some important cases multiple variables distinguished only by subscripts are listed individually.

Variable	Description
$\mathbf{0}$	A vector of 0s.
$\mathbf{1}$	A vector of 1s.
$\vec{a}(\mathbf{r}; \lambda)$	Amplitude spread function of a lens.
$\vec{A}(\mathbf{k}; \lambda)$	Amplitude transfer function of a lens.
B	The indices of the L components present in the object.
b	An arbitrary counting number.
c	The speed of light.
C	The operator matrix excluding spectral envelope effects.
$\mathcal{C}$	The operator matrix including spectral envelope effects.
$d(\mathbf{r})$	The data from an optical system (continuous and noise free).
$\bar{d}(\mathbf{r})$	A continuous representation of sampled data.
$\mathbf{d}$	The noisy discrete data collected from an optical system.
$D(\mathbf{k})$	The Fourier transform of the data
$\bar{D}(\mathbf{k})$	The Fourier transform of the sampled data.
$\bar{\mathbf{D}}(\mathbf{k})$	A vector containing the Fourier-domain data from several channels.
$D_{\text{Comb.}}(\mathbf{k})$	The Fourier-domain result of combining multi-channel data in a linear

system.

$\mathrm{diag}(\cdot)$	An operation that places a vector along the main diagonal of an otherwise zero matrix.
$E[\cdot]$	The expected value operator (a subscript may be added to indicate which variable(s) the expectation is taken over).
$\mathcal{E}$	A matrix that has the envelope vector repeated down its main diagonal.
f	Focal length of a lens.
$\mathcal{F}\{\cdot\}$	The Fourier transform operator.
$g(\mathbf{r}, \mathbf{r}_d; u, v)$	The Green's function relating two spatial points for source polarization u and detection polarization v.
$h(\mathbf{r})$	The point spread function of an optical system.
$H(\mathbf{k})$	The transfer function of an optical system.
$\mathcal{H}(\cdot)$	The Heaviside step function.
I	The identity matrix.
i	$\sqrt{-1}$
$\mathrm{Im}\{\cdot\}$	An operator that returns the imaginary part of the argument.
j, J	The index and limit for number of excitations.
$\mathbf{k}$	The wavenumber.
$\mathbf{k}_c$	A vector containing the cutoff wavenumber in x, y and z directions.
$\mathbf{k}_{\mathrm{Samp.}}$	A vector containing the x, y and z sampling frequencies.
$\mathbf{k}_{\mathrm{Nyq.}}$	A vector containing the x, y and z Nyquist sampling frequencies.
K	An arbitrary constant.
$\mathrm{LPF}[\cdot]$	A low pass filtering operator.
l, L	The index and limit for the number of fluorophore layers.
$\ell, \mathcal{L}$	The index and limit for the number of basis functions in the spline approximation to the envelope function.
m, M	The index and limit for number of points in a measured spectrum.
NA	The numerical aperture of a lens.

n	The index of refraction.
N_x, N_y, N_z	Under-sampling factors in the x, y and z directions.
N	The total under-sampling factor.
$o(\mathbf{r})$	The object to be imaged.
$\mathbf{o}$	A vector representing the discretized object.
$O(\mathbf{k})$	The Fourier transform of the object.
$\mathbf{O}(\mathbf{k})$	A vector containing several points of the object's Fourier transform.
$P(\cdot)$	The probability of the event described by the argument.
p	The type of p-norm.
$\mathcal{P}\{\cdot\,;\cdot\}$	The Poisson distribution where the first argument is the outcome and the second argument is the Poisson parameter.
q, Q	The index and limit for number of points in discretized object.
$\mathrm{Re}\{\cdot\}$	An operator that returns the real part of the argument.
$R(\mathbf{k})$	The pupil filter pattern.
$\mathbf{r}$	The position vector (x, y, z).
$\mathbf{r}_d$	Position in the detector plane.
$\mathbf{r}_o$	The offset between foci in a SSFM system.
$\mathbf{r}_s$	The position of a sample as used for scanning.
$s(\mathbf{r})$	An aperture/pinhole function.
$S(\mathbf{k})$	The Fourier transform of an aperture/pinhole function.
$\mathcal{S}$	The set of spatial points at which sampling occurs.
$SNR(\mathbf{k})$	The signal to noise ratio.
t	Time.
$(\cdot)^T$	The transpose of a matrix.
T	A 3×3 matrix to transform a 3×1 vector.
u, v	Unspecified indices for use in summations, etc.
$w(\mathbf{r})$	A noise function.
$W(\mathbf{k})$	The Fourier transform of the noise function.

x, y	The lateral spatial dimensions.
z	The axial spatial dimension.
α	The regularization parameter.
β	A term defining the continuous-derivative approximation to the p-norm.
$\boldsymbol{\gamma}$	A 2×1 vector defining polarization state.
$\delta(\mathbf{r})$	The Dirac delta function.
δ_{uv}	The Kronecker delta function.
Δ	The integration area used in discretization.
$\epsilon(\lambda)$	The spectral emission envelope of a fluorophore.
$\boldsymbol{\epsilon}$	The discretized spectral envelope vector.
$\boldsymbol{\varepsilon}$	A vector containing the coefficients of the B-spline parameterized envelope.
$\boldsymbol{\zeta}$	A vector of centers of object integration volumes.
$\eta(\mathbf{k})$	A filter used in combining multi-channel data.
θ	The angle from the optic axis of a lens.
Θ	The maximum collected angle from the optic axis in a lens.
$\vartheta^{(\ell)}(\lambda)$	The ℓth spline basis function.
$\kappa(\mathbf{k}, \mathbf{k}^{(u)})$	The spectral purity — a measure of the aliasing effects in an estimate of the object.
λ	The wavelength of light.
$\Lambda(\mathbf{d})$	A matrix with the expected data on the main diagonal. This can also be regarded as an inverse weighting matrix for least squares data fitting.
ν	Temporal optical frequency.
$\boldsymbol{\nu}$	A vector of centers of spectral integration volumes.
$\xi(\mathbf{r})$	The electric field emitted by the object fluorophores at a given position.
Ξ	A matrix that takes B-spline coefficients to the spline curve.
ρ	A reflection coefficient.
ς	The decay coefficient of an evanescent wave

τ	A time delay.
$\Upsilon(\mathbf{o})$	A matrix with the expected data on the main diagonal when spectral envelope effects are not included (c.f. $\Lambda(\mathbf{d})$).
$\Omega(\mathbf{k})$	A matrix containing points from each OTF in a multi-channel microscope.
ϕ	A phase delay.
$\Phi(\mathbf{k}; \alpha)$	The TSVD inverse of $\Omega(\mathbf{k})$ using regularization parameter α.
$\varphi^{(q)}(\mathbf{r})$	The qth component of the object.
$\chi(\mathbf{r})$	The electric field at the detector.
$\Psi(\cdot)$	A cost function (to be minimized).
$\|\cdot\|_p$	The p-norm.
$*$	The convolution operation (subscripts may be added to indicate over which variable(s) this is performed if it is not obvious).
$(\cdot)^*$	The complex conjugate.
$\angle$	The angle of a complex number.

References

Abbe, E. (1873). Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv fur Mikroskopische Anatomie*, 9:413–420.

Agard, D. A., Hiraoka, Y., Shaw, P., and Sedat, J. W. (1989). Fluorescence microscopy in three dimensions. *Methods in Cell Biology*, 30:353–377.

Agard, D. A. and Sedat, J. W. (1983). Three-dimensional architecture of a polytene nucleus. *Nature*, 302:676–681.

Aguet, F., Van De Ville, D., and Unser, M. (2005). A maximum-likelihood formalism for sub-resolution axial localization of fluorescent nanoparticles. *Optics Express*, 13:10503–10522.

Albrecht, B., Failla, A. V., Schweitzer, A., and Cremer, C. (2002). Spatially modulated illumination microscopy allows axial distance resolution in the nanometer range. *Applied Optics*, 41:80–87.

Andresen, V., Egner, A., and Hell, S. W. (2001). Time-multiplexed multifocal multiphoton microscope. *Optics Letters*, 26:75–77.

Andrews, H. C. and Hunt, B. R. (1977). *Digital Image Restoration*. Prentice-Hall.

Arnison, M. R. and Sheppard, C. J. R. (2002). A 3D vectorial optical transfer function suitable for arbitrary pupil functions. *Optics Communications*, 211:53–63.

Axelrod, D., Thompson, N. L., and Burghart, T. P. (1982). Total internal reflection fluorescence microscopy. *Journal of Microscopy*, 129:19–28.

Bahlmann, K. and Hell, S. W. (2000). Polarization effects in 4Pi confocal microscopy studied with water-immersion lenses. *Applied Optics*, 39:1652–1658.

Bailey, B., Farkas, D. L., Taylor, D. L., and Lanni, F. (1993). Enhancement of axial resolution in fluorescence microscopy by standing-wave excitation. *Nature*, 366:44–48.

Barth, M. and Stelzer, E. H. K. (1994). Boosting the optical transfer function with a spatially resolving detector in a high numerical aperture confocal reflection microscope. *Optik*, 96:53–58.

Barton, G. (1989). *Elements of Green's Functions and Propagation*. Oxford University Press.

Basché, T., Moerner, W. E., Orrit, M., and Wild, U. P. (1997). *Single-Molecule Optical Detection, Imaging and Spectroscopy*. VCH.

Begley, R. F., Harvey, A. B., and Byer, R. L. (1974). Coherent anti-Stokes Raman spectroscopy. *Applied Physics Letters*, 25:387–390.

Bertero, M. (1989). *Advances in Electronics and Electron Physics*, volume 75, chapter Linear Inverse and Ill-Posed Problems, pages 1–120. Academic Press.

Bertero, M., Boccacci, P., Brakenhoff, G. J., Malfanti, F., and van der Voort, H. T. M. (1990). Three-dimensional image restoration and super-resolution in fluorescence confocal microscopy. *Journal of Microscopy*, 157:3–20.

Bertero, M., Boccacci, P., Davies, R. E., Malfanti, F., Pike, E. R., and Walker, J. G. (1992). Super-resolution in confocal scanning microscopy: IV. Theory of data inversion by the use of optical masks. *Inverse Problems*, 8:1–23.

Bertero, M., Boccacci, P., Davies, R. E., and Pike, E. R. (1991). Super-resolution in confocal scanning microscopy: III. The case of circular pupils. *Inverse Problems*, 7:655–674.

Bertero, M., Boccacci, P., Defrise, M., De Mol, C., and Pike, E. R. (1989). Super-resolution in confocal scanning microscopy: II. The incoherent case. *Inverse Problems*, 5:441–461.

Bertero, M., Boccacci, P., Malfanti, F., and Pike, E. R. (1994). Super-resolution in confocal scanning microscopy: V. Axial super-resolution in the incoherent case. *Inverse Problems*, 10:1059–1077.

Bertero, M., Boccacci, P., and Pike, E. R. (1982). Resolution in diffraction-limited imaging, a singular value analysis. II. The case of incoherent illumination. *Optica Acta*, 29:1599–1611.

Bertero, M., Brianzi, P., Parker, P., and Pike, E. R. (1984a). Resolution in diffraction-limited imaging, a singular value analysis. III. The effect of sampling and truncation of the data. *Optica Acta*, 31:181–201.

Bertero, M., Brianzi, P., and Pike, E. R. (1987). Super-resolution in confocal scanning microscopy. *Inverse Problems*, 3:195–212.

Bertero, M., De Mol, C., Pike, E. R., and Walker, J. G. (1984b). Resolution in diffraction-limited imaging, a singular value analysis. IV. The case of uncertain localization or non-uniform illumination of the object. *Optica Acta*, 31:923–946.

Bertero, M. and Pike, E. R. (1982). Resolution in diffraction-limited imaging, a singular value analysis. I. The case of coherent illumination. *Optica Acta*, 29:727–746.

Bertsekas, D. P. (1999). *Non-Linear Programming*. Athena Scientific, second edition.

Bewersdorf, J., Pick, R., and Hell, S. W. (1998). Multifocal multiphoton microscopy. *Optics Letters*, 23:655–657.

Biggs, D. S. C. and Andrews, M. (1997). Acceleration of iterative image restoration algorithms. *Applied Optics*, 36:1766–1775.

Blanca, C. M., Bewersdorf, J., and Hell, S. W. (2002). Determination of the unknown phase in 4Pi-confocal microscopy through the image intensity. *Optics Communications*, 206:281–285.

Born, M. and Wolf, E. (1999). *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, seventh edition.

Bracewell, R. (1999). *The Fourier Transform and Its Applications*. McGraw-Hill, third edition.

Braun, D. and Fromherz, P. (1997). Fluorescence interference-contrast microscopy of cell adhension on oxidized silicon. *Applied Physics A Materials Science & Processing*, 65:341–348.

Braun, D. and Fromherz, P. (1998). Fluorescence interferometry of neuronal cell adhension on microstructured silicon. *Physical Review Letters*, 81:5241–5244.

Buist, A. H., Müller, M., Squier, J., and Brakenhoff, G. J. (1998). Real time two-photon absorption microscopy using multi point excitation. *Journal of Microscopy*, 192:217–226.

Carrington, W. A., Fogarty, K. E., Lifschitz, L., and Fay, F. S. (1990). *Handbook of Biological Confocal Microscopy*, chapter Three-dimensional Imaging on Confocal and Wide-field Microscopes, pages 151–161. Plenum Press.

Carrington, W. A., Lynch, R. M., Moore, E. D. W., Isenberg, G., Fogarty, K. E., and Fay, F. S. (1995). Superresolution three-dimensional images of fluorescence in cells with minimal light exposure. *Science*, 268:1483–1487.

Carson, J. R. (1924). A generalization of reciprocal theorem. *Bell Systems Technical Journal*, 3:393–399.

Charbonnier, P., Blanc-Féraud, L., Aubert, G., and Barlaud, M. (1997). Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*, 6:298–311.

Cheung, K. F. (1993). *Advanced Topics in Shannon Sampling and Interpolation Theory*, chapter A Multidimensional Extension of Papoulis' Generalized Sampling Expansion with Applications in Minimum Density Sampling. Springer-Verlag.

Cheung, K. F. and Marks II, R. J. (1985). Ill-posed sampling theorems. *IEEE Transactions on Circuits and Systems*, 32:481–484.

Cheung, K. F. and Marks II, R. J. (1990). Imaging sampling below the Nyquist density without aliasing. *Journal of the Optical Society of America A*, 7:92–105.

Clegg, R. M. (1996). *Fluorescence Imaging Spectroscopy and Microscopy*, volume 137, chapter Fluorescence Resonance Energy Transfer, pages 179–251. John Wiley and Sons.

Coakley, K. J. (1991). A cross-validation procedure for stopping the EM algorithm and deconvolution of neutron depth profiling spectra. *IEEE Transactions on Nuclear Science*, 38:9–15.

Coates, C. G., Denver, D. J., McHale, N. G., Thornbury, K. D., and Hollywood, M. A. (2003). Ultrasensitivity, speed and resolution: Optimizing low-light microscopy with the back-illuminated electron-multiplying CCD. *Proceedings of SPIE*, 5139:56–66. Confocal, Multiphoton, and Nonlinear Microscopic Imaging: T. Wilson: editor.

Conchello, J.-A. and McNally, J. G. (1996). Fast regularization technique for expectation maximization algorithm for optical-sectioning microscopy. *Proceedings of SPIE*, 2655:199–208. Three-Dimensional Microscopy: Image Acquisition and Processing III: C. J. Cogswell, G. S. Kino and T. Wilson: editors.

Cooley, J. W. and Tukey, O. W. (1965). An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19:297–301.

Cox, I. J. and Sheppard, C. J. R. (1986). Information capacity and resolution in an optical system. *Journal of the Optical Society of America A*, 3:1152–1158.

Csiszár, I. (1991). Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. *The Annals of Statistics*, 19:2032–2066.

Davis, B. J., Gough, P., and Hunt, B. (2002). *Impact of Littoral Environmental Variability on Acoustic Predictions and Sonar Performance*, chapter Sea Surface Simulator for Testing a Synthetic Aperture Sonar. Kluwer Academic Publishers.

Davis, B. J., Karl, W. C., Goldberg, B. B., Swan, A. K., and Ünlü, M. S. (2004a). Sampling below the Nyquist rate in interferometric fluorescence microscopy with multi-wavelength measurements to remove aliasing. In *IEEE 11th Digital Signal Processing Workshop and the 3rd Signal Processing Education Workshop*, pages 329–333.

Davis, B. J., Karl, W. C., Goldberg, B. B., Swan, A. K., and Ünlü, M. S. (2005a). Using out-of-focus light to improve image acquisition time in confocal microscopy. *Proceedings of SPIE*, 5701:13–23. Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XII: J.-A. Conchello, C. J. Cogswell and T. Wilson: editors.

Davis, B. J., Karl, W. C., Goldberg, B. B., Swan, A. K., and Ünlü, M. S. (2005b). Using out-of-focus light to improve image acquisition time in confocal microscopy. U.S. Provisional Patent US60/646,776.

Davis, B. J., Karl, W. C., Swan, A. K., Goldberg, B. B., Ünlü, M. S., and Goldberg, M. B. (2004b). Reconstruction of objects with a limited number of non-zero components in fluorescence microscopy. *Proceedings of SPIE*, 5324:27–34. Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XI: J.-A. Conchello, C. J. Cogswell and T. Wilson: editors.

Davis, B. J., Karl, W. C., Swan, A. K., Ünlü, M. S., and Goldberg, B. B. (2004c). Capabilities and limitations of pupil-plane filters for superresolution and image enhancement. *Optics Express*, 12:4150–4156.

Davis, B. J., Kim, E., and Piepmeier, J. R. (2004d). Stochastic modeling and generation of partially polarized or partially coherent electromagnetic waves. *Radio Science*, 39, RS1001.

Davis, B. J. and Nawab, S. H. (2004). The relationship of transform coefficients for differing transforms and/or differing subblock sizes. *IEEE Transactions on Signal Processing*, 52:1458–1461.

Davis, B. J., Ünlü, M. S., Swan, A. K., Goldberg, B. B., and Karl, W. C. (2005c). Using multi-element detectors to create optimal apertures in confocal microscopy. In *The 18th Annual Meeting of the IEEE Lasers and Electro-Optics Society 2005*, pages 547–548.

De Boor, C. (2001). *A Practical Guide to Splines*. Springer.

de Juana, D. M., Oti, J. E., Canales, V. F., and Gagigal, M. P. (2003). Transverse or axial superresolution in a 4Pi-confocal microscope by phase-only filters. *Journal of the Optical Society of America A*, 20:2172–2178.

Defrise, M. and De Mol, C. (1992). Super-resolution in confocal scanning microscopy: Generalized inversion formulae. *Inverse Problems*, 8:175–185.

Delves, L. M. and Mohamed, J. L. (1985). *Computational Methods for Integral Equations*. Cambridge University Press.

Denk, W., Strickler, J. H., and Webb, W. W. (1990). Two-photon laser scanning fluorescence microscopy. *Science*, 248:73–76.

Donoho, D. L. and Huo, X. (2001). Uncertainty principles and ideal atomic decomposition. *IEEE Transactions on Information Theory*, 47:2845–2862.

Donoho, D. L., Johnstone, I. M., Hoch, J. C., and Stern, A. S. (1992). Maximum entropy and the nearly black object. *Journal of the Royal Statistical Society*, 54:41–81.

Drexhage, K. H. (1970). Monomolecular layers and light. *Scientific American*, 6:108–119.

Drexhage, K. H. (1974). Interaction of light with monomolecular dye layers. *Progress in Optics*, 12:163–232.

Duhamel, P. and Vetterli, M. (1990). Fast Fourier transforms: A tutorial review. *Signal Processing*, 19:259–299.

Duncan, M., Reintjes, J., and Manuccia, T. (1982). Scanning coherent anti-Stokes Raman microscope. *Optics Letters*, 7:350–352.

Dyba, M. and Hell, S. W. (2002). Focal spots of size $\lambda/23$ open up far-field fluorescence microscopy at 33nm axial resolution. *Physical Review Letters*, 16.

Egerton, R. F. (2005). *Physical Principles of Electron Microscopy: An Introduction to TEM, SEM, and AEM.* Springer.

Egner, A. and Hell, S. W. (2000). Time multiplexing and parallelization in multifocal multiphoton microscopy. *Journal of the Optical Society of America A*, 17:1192–1200.

Egner, A., Schrader, M., and Hell, S. W. (1998). Refractive index mismatch induced intensity and phase variations in fluorescence confocal, multiphoton and 4Pi-microscopy. *Optics Communications*, 153:211–217.

Elson, E. L. and Madge, D. (1974). Fluorescence correlation spectroscopy. *Biopolymers*, 13:1–27.

Fellgett, P. B. (1951). *The Multiplex Advantage.* PhD thesis, University of Cambridge, U.K.

Fellgett, P. B. (1952). Multi-channel spectrometry. *Journal of the Optical Society of America*, 42:872.

Fittinghoff, D. N. and Squier, J. A. (2000). Time-decorrelated multifocal array for multi-photon microscopy and micromachining. *Optics Letters*, 25:1213–1215.

Frieden, B. R. (1972). Restoring with maximum likelihood and maximum entropy. *Journal of the Optical Society of America*, 62:511–518.

Frohn, J. T., Knapp, H. F., and Stemmer, A. (2001). Three-dimensional resolution enhancement in fluorescence microscopy by harmonic excitation. *Optics Letters*, 26:828–830.

Galatsanos, N. P., Katsaggelos, A. K., Chin, R. T., and Hillery, A. D. (1991). Least squares restoration of multichannel images. *IEEE Transactions on Signal Processing*, 39:2222–2236.

Gannaway, J. and Sheppard, C. J. R. (1978). Second harmonic imaging in the scanning optical microscope. *Optical and Quantum Electronics*, 10:435–439.

Gauderon, R. and Sheppard, C. J. R. (1999). Effect of a finite-size pinhole on noise performance in single-, two-, and three-photon confocal fluorescence microscopy. *Applied Optics*, 38:3562–3565.

Geman, S. and Reynolds, G. (1992). Constrained restoration and the recovery of discontinuities. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:367–383.

Gerchberg, R. W. (1974). Super-resolution through error energy reduction. *Optica Acta*, 21:709–720.

Goldberg, B. B., Swan, A. K., Moiseev, L., Dogan, M., Karl, W. C., Davis, B. J., Cantor, C. R., Ippolito, S. B., Thorne, S. A., Eraslan, M. G., Liu, Z., Goldberg, M. B., and Leblebici, Y. (2004). Seeing inside chips and cells: High-resolution subsurface imaging of integrated circuits, quantum dots and subcellular structures. In *Conference on Lasers and Electro-Optics*, pages 580–581.

Golub, G. H., Heath, M., and Wahba, G. (1979). Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21:215–223.

Golub, G. H. and Van Loan, C. F. (1989). *Matrix Computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press.

Goodman, J. W. (1985). *Statistical Optics*. John Wiley and Sons.

Griffiths, P. R. and De Haseth, J. A. (1986). *Fourier Transform Infrared Spectrometry*. Wiley-Interscience.

Grill, S. and Stelzer, E. H. K. (1999). Method to calculate lateral and axial gain factors of optical setups with a large solid angle. *Journal of the Optical Society of America A*, 16:2658–2665.

Grochmalicki, J., Pike, E. R., and Walker, J. G. (1993). Experimental confirmation of super-resolution in incoherent scanning microscopy. *Pure and Applied Optics: Journal of the European Optical Society Part A*, 2:565–568.

Gu, M. and Sheppard, C. G. R. (1994). Three-dimensional transfer functions in 4Pi confocal microscopes. *Journal of the Optical Society of America A*, 11:1619–1627.

Gu, M. and Sheppard, C. J. R. (1992). Confocal fluorescent microscopy with a finite-sized circular detector. *Journal of the Optical Society of America A*, 9:151–153.

Gu, M. and Sheppard, C. J. R. (1995). Optical transfer function analysis for two-photon 4Pi confocal fluorescence microscopy. *Optics Communications*, 114:45–49.

Gugel, H., Bewersdorf, J., Jakobs, S., Engelhardt, J., Storz, R., and Hell, S. W. (2004). Cooperative 4Pi excitation and detection yields sevenfold sharper optical sections in live-cell microscopy. *Biophysical Journal*, 87:4146–4152.

Gustafsson, M. (2000). Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy. *Journal of Microscopy*, 198:82–87.

Gustafsson, M. G., Agard, D. A., and Sedat, J. W. (1995). Sevenfold improvement of axial resolution in 3D wide-field microscopy using two objective-lenses. *Proceedings of SPIE*, 2412:147–156. Three-Dimensional Microscopy: Image Acquistion and Processing II: T. Wilson and C. J. Cogswell: editors.

Gustafsson, M. G. L. (1999). Extended resolution fluorescence microscopy. *Current Opinion in Structural Biology*, 9:627–634.

Gustafsson, M. G. L. (2005). Nonlinear structured-illumination microscopy: Wide-field fluorescence imaging with theoretically unlimited resolution. *Proceedings of the National Academy of Sciences*, 102:13081–13086.

Gustafsson, M. G. L., Agard, D. A., and Sedat, J. W. (1999). I^5M: 3D widefield light microscopy with better than 100nm axial resolution. *Journal of Microscopy*, 195.

Hanisch, R. J., White, R. L., and Gilliland, R. L. (1997). *Deconvolution of Images and Spectra*, chapter Deconvolution of Hubble Space Telescope Images and Spectra, pages 310–360. Academic Press, second edition.

Hansen, P. C. (1992). Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Review*, 34:561–580.

Hansen, P. C. (1998). *Rank-Deficient and Discrete Ill-Posed Problems*. SIAM.

Harikumar, G. and Bresler, Y. (1999). Exact image deconvolution from multiple FIR blurs. *IEEE Transactions on Image Processing*, 8:846–862.

Heintzmann, R. and Cremer, C. (2002). Axial tomographic confocal fluorescence microscopy. *Journal of Microscopy*, 206:7–23.

Heintzmann, R. and Jovin, T. M. (2002). Saturated patterned excitation microscopy — a concept for optical resolution improvement. *Journal of the Optical Society of America*, 19:1599–1609.

Heintzmann, R., Sarafis, V., Munroe, P., Nailon, J., Hanley, Q. S., and Jovin, T. M. (2003). Resolution enhancement by subtraction of confocal signals taken at different pinhole sizes. *Micron*, 34:293–300.

Hell, S. and Stelzer, E. H. K. (1992a). Properties of a 4Pi confocal fluorescence microscope. *Journal of the Optical Society of America A*, 9:2159–2166.

Hell, S. W. (1990). Double-scanning microscope. European patent 0491289.

Hell, S. W. (2003). Toward fluorescence nanoscopy. *Nature Biotechnology*, 21:1347–1355.

Hell, S. W., Blanca, C. M., and Bewersdorf, J. (2002). Phase determination in interference-based superresolving microscopes through critical frequency analysis. *Optics Letters*, 27:888–890.

Hell, S. W. and Nagorni, M. (1998). 4Pi confocal microscopy with alternate interference. *Optics Letters*, 23:1567–1569.

Hell, S. W., Schrader, M., and van der Voort, H. T. M. (1997). Far-field fluorescence microscopy with three-dimensional resolution in the 100-nm range. *Journal of Microscopy*, 187:1–7.

Hell, S. W. and Stelzer, E. H. K. (1992b). Fundamental improvement of resolution with a 4Pi-confocal fluorescence microscope using two-photon excitation. *Optics Communications*, 93:277–282.

Hell, S. W. and Wichmann, J. (1994). Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Optics Letters*, 19:780–782.

Holmes, T. J. (1988). Maximum-likelihood image restoration adapted for noncoherent optical imaging. *Journal of the Optical Society of America A*, 5:666–673.

Holmes, T. J. (1989). Expectation-maximization restoration of band-limited, truncated point-process intensities with application in microscopy. *Journal of the Optical Society of America A*, 6:1006–1014.

Holmes, T. J. (1992). Blind deconvolution of quantum-limited incoherent imagery: Maximum-likelihood approach. *Journal of the Optical Society of America A*, 9:1052–1061.

Holmes, T. J. and O'Connor, N. J. (2000). Blind deconvolution of 3D transmitted light brightfield micrographs. *Journal of Microscopy*, 200:114–127.

Homem, M. R. P., Mascarenhas, N. D. A., Costa, L. F., and Preza, C. (2002). Biological image restoration in optical-sectioning microscopy using prototype image constraints. *Real Time Imaging*, 8:475–490.

Huang, D., Swanson, E. A., Lin, C. P., Schuman, J. S., Stinson, W. G., Chang, W., Hee, M. R., Flotte, T., Gregory, K., Puliafito, C. A., and Fujimoto, J. G. (1991). Optical coherence tomography. *Science*, 254:1178–1181.

Huisken, J., Stelzer, E. H. K., and Swoger, J. (2005). Selective plane illumination microscopy: Optical sections from many directions. *G. I. T. Imaging and Microscopy*, 1:2–3.

Huisken, J., Swoger, J., Del Bene, F., Wittbrodt, J., and Stelzer, E. H. K. (2004). Optical sectioning deep inside live embryos by selective plane illumination microscopy. *Science*, 305:1007–1009.

Jain, A. K. (1988). *Fundamentals of Digital Image Processing.* Prentice Hall.

Kano, H., Jakobs, S., Nagorni, M., and Hell, S. W. (2002). Dual-color 4Pi-confocal microscopy with 3D-resolution in the 100nm range. *Ultramicroscopy*, 90:207–213.

Karl, W. C. (2000). *Handbook of Image and Video Processing*, chapter Regularization in Image Restoration and Reconstruction, pages 141–161. Academic Press.

Kay, S. M. (1993). *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall.

Klar, T. A., Engel, E., and Hell, S. W. (2001). Breaking Abbe's diffraction resolution limit in fluorescence microscopy with stimulated emission depletion beams of various shapes. *Physical Review E*, 64.

Klar, T. A. and Hell, S. W. (1999). Subdiffraction resolution in far-field fluorescence microscopy. *Optics Letters*, 24:954–956.

Kong, J. A. (1990). *Electromagnetic Wave Theory.* Wiley, New York.

Kress, R. (1989). *Linear Integral Equations.* Springer-Verlag.

Krishnamurthi, V., Bailey, B., and Lanni, F. (1996). Image processing in 3D standing-wave fluorescence microscopy. *Proceedings of SPIE*, 2655:18–25. Three-Dimensional Microscopy: Image Acquisition and Processing III: C. J. Cogswell, G. S. Kino and T. Wilson: editors.

Lakowicz, J. R., Szmacinski, H., Nowaczyk, K., Berndt, K. W., and Johnson, M. (1992). Fluorescence lifetime imaging. *Analytical Biochemistry*, 202:316–330.

Lambacher, A. and Fromherz, P. (1996). Fluorescence interference-contrast microscopy on oxidized silicon using a monomolecular dye layer. *Applied Physics A Materials Science & Processing*, 63:207–216.

Landon, L., Lifshitz, E. M., and Pitaevskii, L. P. (1984). *Electrodynamics of Continuous Media*. Butterworth-Heinemann.

Liebman, P. A. and Entine, G. (1974). Lateral diffusion of visual pigment in photoreceptor disk membranes. *Science*, 185:457–459.

Lorentz, H. A. (1896). The theorem of Poynting concerning the energy in the electromagnetic field and two general propositions concerning the propagation of light. *Amsterdammer Akademie der Wetenschappen*, 4:176.

Lucy, L. B. (1974). An iterative technique for the rectification of observed distributions. *The Astronomical Journal*, 79:745–754.

Madge, D., Elson, E. L., and Webb, W. W. (1974). Fluorescence correlation spectroscopy. II. Experimental realization. *Biopolymers*, 13:29–61.

Malioutov, D. (2003). A sparse signal reconstruction perspective for source localization with sensor arrays. Master's thesis, Massachusetts Institute of Technology.

Markham, J. and Conchello, J. (2001). Fast maximum-likelihood image restoration algorithms for three-dimensional fluorescence microscopy. *Journal of the Optical Society of America A*, 18:1062–1071.

Martínez-Corral, M., Caballero, M. T., Stelzer, E. H. K., and Swoger, J. (2002). Tailoring the axial shape of the point spread function using the Toraldo concept. *Optics Express*, 10:98–103.

Martínez-Corral, M., Ibáñez-López, C., and Saavedra, G. (2003). Axial gain resolution in optical sectioning fluorescence microscopy by shaded-ring filters. *Optics Express*, 11:1740–1745.

McCutchen, C. W. (1964). Generalized aperture and the three-dimensional diffraction image. *Journal of the Optical Society of America*, 54:240–244.

McCutchen, C. W. (1991). Convolution relation within the three-dimensional diffraction image. *Journal of the Optical Society of America A*, 8:868–870.

McCutchen, C. W. (2002). Generalized aperture and the three-dimensional diffraction image: Erratum. *Journal of the Optical Society of America A*, 19:1781.

Meyvis, T. K., De Smedt, S. C., Van Oostveldt, P., and Demeester, J. (1999). Fluorescence recovery after photobleaching: A versatile tool for mobility and interaction measurements in pharmaceutical research. *Pharmaceutical Research*, 16:1153–1162.

Minsky, M. (1955). Confocal scanning microscope. U.S. Patent 3013467.

Moiseev, L., Cantor, C. R., Aksun, M. I., Dogan, M., Goldberg, B. B., Swan, A. K., and Ünlü, M. S. (2004a). Spectral self-interference fluorescence microscopy. *Journal of Applied Physics*, 96:5311–5315.

Moiseev, L., Cantor, C. R., Swan, A. K., Goldberg, B. B., and Ünlü, M. S. (2004b). Biological applications of spectral self-interference. *Proceedings of SPIE*, 5331:36–43. Nanobiophotonics and Biomedical Applications: Alexander N. Cartwright: editor.

Moiseev, L., Swan, A. K., Ünlü, M. S., Goldberg, B. B., and Cantor, C. R. (2004c). Spectral self-interference on fluorescently labeled DNA monolayers. In *Proceedings of the IEEE Lasers and Electro-Optics Society 2004 Annual Meeting*.

Moiseev, L., Ünlü, M. S., Swan, A. K., Goldberg, B. B., and Cantor, C. R. (2006). DNA conformation on surfaces measured by fluorescence self-interference. *Proceedings of the National Academy of Sciences*, 103:2623–2628.

Mugnai, D., Ranfagni, A., and Ruggeri, R. (2003). Pupils with super-resolution. *Physics Letters A*, 311:77–81.

Murphy, D. B. (2001). *Fundamentals of Light Microscopy and Electronic Imaging*. Wiley-Liss.

Murray, J. M. (1998). Evaluating the performance of fluorescence microscopes. *Journal of Microscopy*, 191:128–134.

Nagorni, M. and Hell, S. W. (1998). 4Pi-confocal microscopy provides three-dimensional images of the microtubule network with 100- to 150-nm resolution. *Journal of Structural Biology*, 123:236–247.

Nagorni, M. and Hell, S. W. (2001a). Coherent use of opposing lenses for axial resolution increase. II. Power and limitation of nonlinear image restoration. *Journal of the Optical Society of America A*, 18:49–54.

Nagorni, M. and Hell, S. W. (2001b). Coherent use of opposing lenses for axial resolution increase in fluorescence microscopy. I. Comparitive study of concepts. *Journal of the Optical Society of America A*, 18:36–48.

Neil, M. A. A., Juškaitis, R., Wilson, T., and Laczik, Z. J. (2000). Optimized pupil-plane filters for confocal microscope point-spread function engineering. *Optics Letters*, 25:245–247.

Neil, M. A. A., Massoumian, F., Juškaitis, R., and Wilson, T. (2002). Method for the generation of arbitrary complex vector wave fronts. *Optics Letters*, 27:1929–1931.

Nomarski, G. (1955). Microinterféromètre différentiel à ondes polarisées. *Journal de Physique et le Radium*, 16:S9–S13.

Ober, R. J., Ram, S., and Ward, E. S. (2004). Localization accuracy in single-molecule microscopy. *Biophysical Journal*, 86:1185–1200.

Oliver, C. J. and Pike, E. R. (1974). Multiplex advantage in the detection of optical images in the photon noise limit. *Applied Optics*, 13:158–161.

Oppenheim, A. V. and Schafer, R. W. (1989). *Discrete-Time Signal Processing*. Prentice Hall.

Papoulis, A. (1975). A new algorithm in spectral analysis and band-limited extrapolation. *IEEE Transactions on Circuits and Systems*, CAS-22:735–742.

Papoulis, A. (1977). Generalized sampling expansion. *IEEE Transactions on Circuits and Systems*, CAS-24(11):652–654.

Papoulis, A. (1984). *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Series in Electrical Engineering. McGraw-Hill.

Parker, J. R. (1996). *Algorithms for Image Processing and Computer Vision*. Wiley.

Pawley, J. (1994). *Three Dimensional Confocal Microscopy: Volume Investigation of Biological Specimens*, chapter Sources of Noise in Three-Dimensional Microscopical Data Sets, pages 47–94. Academic Press.

Pawley, J. (2000). The 39 steps. A cautionary tale of quantitative 3D fluorescence microscopy. *BioTechniques*, 28:884–888.

Pawley, J., Blouke, M., and Janesick, J. (1996). The CCDiode: An optimal detector for laser confocal microscopes. *Proceedings of SPIE*, 2655:125–129. Three-Dimensional Microscopy: Image Acquisition and Processing III: C. J. Cogswell, G. S. Kino and T. Wilson: editors.

Periasamy, A., editor (2001). *Methods in Cellular Imaging*. An American Physiological Society Book.

Petráň, M., Hadravský, M., Egger, M. D., and Galambos, R. (1968). Tandem scanning optical microscope. *Journal of the Optical Society of America*, 58:661–664.

Petty, H. R. (2004). Applications of high speed microscopy in biomedical research. *Optics and Photonics News*, 15:40–45.

Pohl, D. W. and Courjon, D. (1993). *Near Field Optics*. Kluwer.

Potton, R. J. (2004). Reciprocity in optics. *Reports on Progress in Physics*, 67:717–754.

Ram, S., Ward, E. S., and Ober, R. J. (2005). How accurately can a single molecule be localized in three dimensions using a fluorescence microscope? *Proceedings of SPIE*, 5699:426–435. Imaging, Manipulation, and Analysis of Biomolecules and Cells: Fundamentals and Applications III: D. V. Nicolau, R. C. Leif, D. L. Farkas and R. Raghavachari: editors.

Ram, S., Ward, E. S., and Ober, R. J. (2006). A stochastic analysis of performance limits for optical microscopes. *Multidimensional Systems and Signal Processing*, 17:27–57.

Richards, B. and Wolf, E. (1959). Electromagnetic diffraction in optical systems. II. Structure of the image field in an aplanatic system. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 253:358–379.

Richardson, W. H. (1972). Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62:55–59.

Saleh, B. E. A. and Teich, M. C. (1991). *Fundamentals of Photonics*, chapter Fourier Optics, pages 136–139. Wiley-Interscience.

Schmidt, M., Nagorni, M., and Hell, S. W. (2000). Subresolution axial distance measurements in far-field fluorescence microscopy with precision of 1 nanometer. *Review of Scientific Instruments*, 71:2742–2745.

Schmitt, J. M. (1999). Optical coherence tomography (OCT): A review. *IEEE Journal of Selected Topics in Quantum Electronics*, 5:1205–1215.

Schönle, A. and Hell, S. W. (2002). Calculation of vectorial three-dimensional transfer functions in large-angle focusing systems. *Journal of the Optical Society of America A*, 19:2121–2126.

Schrader, M., Bahlmann, K., Giese, G., and Hell, S. W. (1998a). 4Pi-confocal imaging in fixed biological specimens. *Biophysical Journal*, 75:1659–1668.

Schrader, M., Hell, S. W., and van der Voort, H. T. M. (1998b). Three-dimensional super-resolution with a 4Pi-confocal microscope using image restoration. *Journal of Applied Physics*, 84:4033–4042.

Schrader, M., Kozubek, M., Hell, S. W., and Wilson, T. (1997). Optical transfer functions of 4Pi confocal microscopes: Theory and experiment. *Optics Letters*, 22:436–438.

Sheppard, C. J. R. (1986a). The spatial frequency cut-off in three-dimensional imaging. *Optik*, 72:131–133.

Sheppard, C. J. R. (1986b). The spatial frequency cut-off in three-dimensional imaging II. *Optik*, 74:128–129.

Sheppard, C. J. R. (1995). Leaky annular pupils for improved axial imaging. *Optik*, 99:32–34.

Sheppard, C. J. R. and Cogswell, C. J. (1990). Confocal microscopy with detector arrays. *Journal of Modern Optics*, 37:267–279.

Sheppard, C. J. R., Cogswell, C. J., and Gu, M. (1991). Signal strength and noise in confocal microscopy: Factors influencing selection of an optimum detector aperture. *Scanning*, 13:233–240.

Sheppard, C. J. R. and Gu, M. (1994). Three-dimensional transfer functions for high-aperture systems. *Journal of the Optical Society of America A*, 11:593–598.

Sheppard, C. J. R., Gu, M., and Roy, M. (1992). Signal-to-noise ratio in confocal microscope systems. *Journal of Microscopy*, 168:209–218.

Sheppard, C. J. R. and Hegedus, Z. S. (1988). Axial behavior of pupil-plane filters. *Journal of the Optical Society of America A*, 5:643–747.

Sheppard, C. J. R. and Matthews, H. J. (1987). Imaging in high-aperture optical systems. *Journal of the Optical Society of America A*, 4:1354–1360.

Sheppard, C. J. R. and Wilson, T. (1981). The halo effect of image processing by spatial frequency filtering. *Optik*, 56:19–23.

Shinohara, K. and Yada, K. (1990). *X-Ray Microscopy in Biology and Medicine*. Springer.

Shotton, D. M., editor (1993). *Electronic Light Microscopy: The Principles and Practice of Video-Enhanced Contrast, Digital Intensified Fluorescence, and Confocal Scanning Light Microscopy*. Techniques in Modern Biomedical Microscopy. Wiley-Liss.

Slayter, E. M. and Slayter, H. S. (1992). *Light and Electron Microscopy*. Cambridge University Press.

Smith, F. H. (1955). Microscopic interferometry. *Research (London)*, 8:385–395.

Snyder, D. L., Schulz, T., and O'Sullivan, J. A. (1992). Deblurring subject to nonnegativity constraints. *IEEE Transactions on Signal Processing*, 40:1143–1150.

Stelzer, E. H. K. (1998). Contrast, resolution, pixelation, dynamic range and signal-to-noise ratio: Fundamental limits to resolution in fluorescence light microscopy. *Journal of Microscopy*, 189:15–24.

Stelzer, E. H. K. and Lindek, S. (1994). Fundamental reduction of the observation volume in far-field light microscopy by detection orthogonal to the illumination axis: Confocal theta microscopy. *Optics Communications*, 111:536–547.

Strang, G. (1986). *Introduction to Applied Mathematics*. Wellesley-Cambridge Press.

Stryer, L. (1978). Fluorescence energy transfer as a spectroscopic ruler. *Annual Review of Biochemistry*, 47:819–846.

Swan, A. K., Moiseev, L., Tong, Y., Lipoff, S. H., Karl, W. C., Goldberg, B. B., and Ünlü, M. S. (2002). High resolution spectral self-interference fluorescence microscopy. *Proceedings of SPIE*, 4621:77–85. Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing IX: J.-A. Conchello, C. J. Cogswell and T. Wilson: editors.

Swan, A. K., Moiseev, L. A., Cantor, C. R., Davis, B., Ippolito, S. B., Karl, W. C., Goldberg, B. B., and Ünlü, M. S. (2003). Toward nanometer-scale resolution in fluorescence microscopy using spectral self-interference. *IEEE Journal of Selected Topics in Quantum Electronics*, 9:294–300.

Swedlow, J. R., Sedat, J. W., and Agard, D. A. (1997). *Deconvolution of Images and Spectra*, chapter Deconvolution in Optical Microscopy, pages 284–309. Academic Press, second edition.

Swoger, J., Huisken, J., and Stelzer, E. H. K. (2003). Multiple imaging axis microscopy improves resolution for thick-sample applications. *Optics Letters*, 28:1654–1656.

Thompson, N. L. (1991). *Topics in Fluorescence Spectroscopy*, chapter Fluorescence correlation spectroscopy, pages 337–378. Plenum Press.

Tikhonov, A. N. (1963a). Regularization of incorrectly posed problems. *Soviet Mathematics Doklady*, 4:1624–1627.

Tikhonov, A. N. (1963b). Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics Doklady*, 4:1035–1038.

Toraldo di Francia, G. (1952). Nuovo pupille superresolventi. *Atti della Fondazione Giorgio Ronchi*, 7:366–372.

Török, P. and Kao, F.-J., editors (2003). *Optical Imaging and Microscopy: Techniques and Advanced Systems*. Springer.

Tsien, R. Y. and Waggoner, A. (1990). *Handbook of Biological Confocal Microscopy*, chapter Fluorophores for Confocal Microscopy: Photophysics and Photochemistry, pages 169–178. Plenum Press.

Ünlü, M. S. and Davis, B. J. (2005). Multi-element photodetector as a smart pinhole in confocal microscopy. U.S. Provisional Patent US60/669,602.

Vaidyanathan, P. P. (1992). *Multirate Systems and Filter Banks*. Prentice Hall PTR.

van de Nes, A. S., Billy, L., Pereira, S. F., and Braat, J. J. M. (2004). Calculation of the vectorial field distribution in a stratified focal region of a high numerical aperture imaging system. *Optics Express*, 12:1281–1293.

van der Voort, H. T. M. and Brakenhoff, G. J. (1990). 3-D image formation in high-aperture fluorescence confocal microscopy: A numerical analysis. *Journal of Microscopy*, 158:43–54.

van Kempen, G. M. P., van Vliet, L. J., Verveer, P. J., and van der Voort, H. T. M. (1997). A quantitative comparison of image restoration methods for confocal microscopy. *Journal of Microscopy*, 185:354–365.

Venkataramani, R. and Bresler, Y. (2003). Sampling theorems for uniform and periodic nonuniform MIMO sampling of multiband signals. *IEEE Transactions on Signal Processing*, 51:3152–3163.

Verveer, P. J., Gemkow, M. J., and Jovin, T. M. (1999). A comparison of image restoration approaches applied to three-dimensional confocal and wide-field fluorescence microscopy. *Journal of Microscopy*, 193:50–61.

Verveer, P. J., Hanley, Q. S., Verbeek, P. W., Van Vliet, L. J., and Jovin, T. M. (1998). Theory of confocal fluorescence imaging in the programmable array microscope PAM. *Journal of Microscopy*, 189:192–198.

Verveer, P. J. and Jovin, T. M. (1998). Improved restoration from multiple images of a single object: Application to fluorescence microscopy. *Applied Optics*, 37:6240–6246.

Walker, J. G. (1983). Optical imaging with resolution exceeding the Rayleigh criterion. *Optica Acta*, 30:1197–1202.

Walker, J. G., Pike, E. R., Davies, R. E., Young, M. R., Brakenhoff, G. J., and Bertero, M. (1993). Superresolving scanning optical microscopy using holographic optical processing. *Journal of the Optical Society of America A*, 10:59–64.

Wallace, W., Schaefer, L. H., and Swedlow, J. R. (2001). A workingperson's guide to deconvolution in light microscopy. *BioTechniques*, 31:1076–1097.

Webb, R. H. and Dorey, C. K. (1990). *Handbook of Biological Confocal Microscopy*, chapter The Pixelated Image, pages 41–51. Plenum Press.

White, J. and Stelzer, E. (1999). Photobleaching GFP reveals protein dynamics inside live cells. *Trends in Cell Biology*, 9:61–65.

Wiener, N. (1964). *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. The MIT Press.

Wilson, T. (1990). *Handbook of Biological Confocal Microscopy*, chapter The Role of the Pinhole in Confocal Imaging Systems, pages 113–126. Plenum Press.

Wilson, T. and Sheppard, C. J. R. (1984). *Theory and Practice of Scanning Optical Microscopy*. Academic Press.

Wohlberg, B. (2003). Noise sensitivity of sparse signal representations: Reconstruction error bounds for the inverse problem. *IEEE Transactions on Signal Processing*, 51:3053–3060.

Wolf, E. (1959). Electromagnetic diffraction in optical systems. I. An integral representation of the image field. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 253:349–357.

Worth, C., Goldberg, B. B., Ruane, M., and Ünlü, M. S. (2001). Surface desensitization of polarimetric waveguide interferometers. *IEEE Journal on Selected Topics in Quantum Electronics*, 7:874–877.

Xiao, G. Q., Corle, T. R., and Kino, G. S. (1988). Real-time confocal scanning optical microscope. *Applied Physics Letters*, 53:716–718.

Youngworth, K. and Brown, T. (2000). Focusing of high numerical aperture cylindrical-vector beams. *Optics Express*, 7:77–87.

Zernike, F. (1955). How I discovered phase contrast. *Science*, 121:345–349.

Zumbusch, A., Holtom, G. R., and Xie, X. S. (1999). Vibrational microscopy using coherent anti-Stokes Raman scattering. *Physical Review Letters*, 82:4142–4145.

CURRICULUM VITAE

Brynmor Davis

RESEARCH INTERESTS

Inverse Problems	Statistical Signal Processing	Microscopy
Optics and Acoustics	Computer Vision and Image Processing	Sensing and Imaging

EDUCATION

PhD — Boston University, Boston, MA   May 2006
  · Electrical and Computer Engineering
  · 4.0 Grade Point Average

M.S. — The University of Arizona, Tucson, AZ   Jul 2001
  · Electrical and Computer Engineering with an Optical Sciences minor
  · 4.0 Grade Point Average

B.E. — The University of Canterbury, Christchurch, New Zealand   Dec 1998
  · Electrical and Electronic Engineering
  · First Class Honors

GRADUATE COURSES

Recursive Estimation and Optimal Filtering	Quantum Optics
Detection and Estimation in Engineering Systems	Computer Vision
Random Processes for Engineering Applications	Diffraction and Interferometry
Analytical Methods in Electrical Engineering	Digital Image Processing
Introduction to Fourier Optics	Medical Optics
Image Reconstruction and Restoration	Statistical Pattern Recognition
Advanced Digital Signal Processing	Discrete Stochastic Models
Non-linear and Ultra-fast Optics	Digital Signal Processing
Information Theory and Coding	Error-Control Codes
Subsurface Sensing and Imaging	Advanced Discrete Mathematics

EXPERIENCE

Boston University, Boston, MA   May 2002 – Mar 2006
Research Assistant, Fluorescence Microscopy
    · Designed data inversion algorithms
    · Developed physical models
    · Predicted instrument performance in different operating regimes and in comparison
      to other systems

German Cancer Research Center, Heidelberg, Germany   Feb 2005 – May 2005
Visiting Researcher, 4Pi Microscopy
    · Developed deconvolution algorithms for a two-channel microscopy system
    · Numerically estimated point spread functions from data

The University of Arizona, Tucson, AZ   Aug 1999 – Jun 2001
Research Assistant, Synthetic Aperture Sonar
    · Characterized multi-path propagation effects

NASA Goddard Space Flight Center, Greenbelt, MD　　　Jun 2000 – Aug 2000
Visiting Student Researcher, Correlation Radiometers
　　· Extended signal modeling theory
　　· Developed novel calibration techniques
The Australian National University, Canberra, Australia　　Dec 1998 – Feb 1999
Summer Research Scholar, Fading Channel Communications
　　· Simulated Maximum Likelihood Sequence Detection (MLSD) over a time-varying
　　dispersive channel
Mainpower New Zealand Ltd., Rangiora, New Zealand　　Nov 1997 – Feb 1998
Engineer's Assistant, Power Distribution Networks
　　· Power distribution network simulation

COMPUTER SKILLS

MATLAB　　　　　　　　　　LaTeX　　　　　　　　　　Microsoft Office

PATENTS

· US Provisional Patent US60/646,776, "Using out-of-focus light to improve image acquisition time in confocal microscopy," Brynmor J. Davis, William C. Karl, Bennett B. Goldberg, Anna K. Swan, M. Selim Ünlü, 2005
· US Provisional Patent US60/646,776, "Multi-element photodetector as a smart pinhole in confocal microscopy," M. Selim Ünlü, Brynmor J. Davis, 2005

SELECTED PUBLICATIONS

· "Capabilities and limitations of pupil-plane filters for superresolution and image enhancement," Brynmor J. Davis, William C. Karl, Anna K. Swan, M. Selim Ünlü and Bennett B. Goldberg, Optics Express, Vol. 12, No. 17, August 2004, pp 4150-4156
· "The relationship of transform coefficients for differing transforms and/or differing sub-block sizes," Brynmor J. Davis and S. Hamid Nawab, IEEE Transactions on Signal Processing, Vol. 52, No. 5, May 2004, pp 1458-1461
· "Stochastic modeling and generation of partially polarized or partially coherent electromagnetic waves," Brynmor Davis, Edward Kim and J. R. Piepmeier, Radio Science, Vol. 39, RS1001, January 2004
· "Toward nanometer-scale resolution in fluorescence microscopy using spectral self-interference," A. K. Swan, L. A. Moiseev, C. R. Cantor, B. Davis, S. B. Ippolito, W. C. Karl, B. B. Goldberg and M. S. Ünlü, IEEE Journal of Selected Topics in Quantum Electronics, Vol. 9, No. 2, March/April 2003, pp 294-300

ACHIEVEMENTS

· Awarded 2004 IEEE 11th Digital Signal Processing Workshop student fellowship
· Member of 'Boston Miracles' men's field hockey team — North American Premier League (NAPL) champions in 2002 and 2003
· Recipient of Boston University Presidential University Graduate Fellowship in 2001
· Achieved direct entry into sophomore year of the University of Canterbury's engineering program based on high school grades