Blind Deconvolution for Confocal Laser Scanning Microscopy

Praveen Pankajakshan

To cite this version:

Praveen Pankajakshan. Blind Deconvolution for Confocal Laser Scanning Microscopy. Signal and Image processing. Université Nice Sophia Antipolis, 2009. English. NNT : . tel-00474264

HAL Id: tel-00474264

https://theses.hal.science/tel-00474264v1

Submitted on 19 Apr 2010

UNIVERSITY OF NICE - SOPHIA ANTIPOLIS

DOCTORAL SCHOOL STIC
INFORMATION AND COMMUNICATION SCIENCES AND TECHNOLOGIES

THESIS

to fulfill the requirements for the degree of
Doctor of Philosophy in Computer Science
from the University of Nice - Sophia Antipolis
Specialized in Control, Signal and Image Processing
by
Praveen PANKAJAKSHAN

Blind Deconvolution for Confocal Laser Scanning Microscopy

Supervised by Laure BLANC-FÉRAUD

and prepared at INRIA Sophia Antipolis Méditerranée in the Ariana research team,
defended on 15 December 2009, in front of the committee composed of

Laure Blanc-Féraud,	Research director, CNRS	-	Advisor
Philippe Ciuciu,	Full-time researcher, CEA	-	Reviewer
Alain Dieterlen,	Professor, IUT of Mulhouse	-	Reviewer
Gilbert Engler,	Responsible of the microscopy platform, INRA	-	Member
Zvi Kam,	Professor, Weizmann Institute of Sciences	-	Member
Jean-Christophe Olivo-Marin,	Research director, Pasteur Institute	-	Member
James B. Pawley,	Professor, University of Wisconsin-Madison	-	President
Josiane Zerubia,	Research director, INRIA	-	Co-advisor

UNIVERSITÉ DE NICE - SOPHIA ANTIPOLIS

ECOLE DOCTORALE STIC
SCIENCES ET TECHNOLOGIES DE L'INFORMATION ET DE LA COMMUNICATION

THÈSE

pour obtenir le titre de
Docteur en Sciences
de l'Université de Nice - Sophia Antipolis
Mention : Automatique, traitement du signal et des images
par
Praveen PANKAJAKSHAN

Déconvolution Aveugle en Imagerie de Microscopie Confocale À Balayage Laser

---

Thèse dirigée par Laure BLANC-FÉRAUD

et préparée à l'INRIA Sophia Antipolis Méditerranée dans le projet Ariana
Soutenue le 15 Decembre 2009 devant le jury suivant :

Laure Blanc-Féraud,	Directrice de recherche, CNRS	-	Directrice
Philippe Ciuciu,	Ingénieur de recherche, CEA	-	Rapporteur
Alain Dieterlen,	Professeur, IUT de Mulhouse	-	Rapporteur
Gilbert Engler,	Responsable du plateau de microscopie, INRA	-	Membre
Zvi Kam,	Professeur, Institut des sciences Weizmann	-	Membre
Jean-Christophe Olivo-Marin,	Directeur de recherche, Institut Pasteur Paris	-	Membre
James B. Pawley,	Professeur, Université du Wisconsin-Madison	-	President
Josiane Zerubia,	Directrice de recherche, INRIA	-	Co-directrice

I, Praveen Pankajakshan, confirm that the work presented in this thesis is my own. Where information was derived from other sources, I confirm that this is amply referenced to the best of my knowledge.

Abstract

Confocal laser scanning microscopy is a powerful technique for studying biological specimens in three dimensions (3D) by optical sectioning. It permits to visualize images of live specimens non-invasively with a resolution of few hundred nanometers. Although ubiquitous, there are uncertainties in the observation process. As the system's impulse response, or point-spread function (PSF), is dependent on both the specimen and imaging conditions, it should be estimated from the observed images in addition to the specimen. This problem is ill-posed, under-determined. To obtain a solution, it is necessary to insert some knowledge in the form of *a priori* and adopt a Bayesian approach. The state of the art deconvolution and blind deconvolution algorithms are reviewed within a Bayesian framework. In the first part, we recognize that the diffraction-limited nature of the objective lens and the intrinsic noise are the primary distortions that affect specimen images. An alternative minimization (AM) approach restores the lost frequencies beyond the diffraction limit by using total variation regularization on the object, and a spatial constraint on the PSF. Additionally, some methods are proposed to ensure positivity of estimated intensities, to conserve the object's flux, and to well handle the regularization parameter. When imaging thick specimens, the phase of the pupil function due to spherical aberration (SA) cannot be ignored. It is shown to be dependent on the refractive index mismatch between the object and the objective immersion medium, and the depth under the cover slip. The imaging parameters and the object's original intensity distribution are recovered by modifying the AM algorithm. Due to the incoherent nature of the light in fluorescence microscopy, it is possible to retrieve the phase from the observed intensities by using a model derived from geometrical optics. This was verified on the simulated data. This method could also be extended to restore specimens affected by SA. As the PSF is space varying, a piecewise convolution model is proposed, and the PSF approximated so that, apart from the specimen, it is sufficient to estimated only one free parameter.

Keywords: confocal laser scanning microscopy, point-spread function, blind deconvolution, Bayesian approach, maximum likelihood, EM algorithm, total variation, maximum *a posteriori*, alternate minimization, spherical aberration.

Resumé

La microscopie confocale à balayage laser, est une technique puissante pour étudier les spécimens biologiques en trois dimensions (3D) par sectionnement optique. Elle permet d'avoir des images de spécimen vivants à une résolution de l'ordre de quelques centaines de nanomètres. Bien que très utilisée, il persiste des incertitudes dans le procédé d'observation. Comme la réponse du système à une impulsion, ou fonction de flou (PSF), est dépendante à la fois du spécimen et des conditions d'acquisition, elle devrait être estimée à partir des images observées du spécimen. Ce problème est mal posé et sous déterminé. Pour obtenir une solution, il faut injecter des connaisances, c'est à dire, *a priori* dans le problème. Pour cela, nous adoptons une approche bayésienne. L'état de l'art des algorithmes concernant la déconvolution et la déconvolution aveugle est exposé dans le cadre d'un travail bayésien. Dans la première partie, nous constatons que la diffraction due à l'objectif et au bruit intrinsèque à l'acquisition, sont les distorsions principales qui affectent les images d'un spécimen. Une approche de minimisation alternée (AM), restaure les fréquences manquantes au-delà de la limite de diffraction, en utilisant une régularisation par la variation totale sur l'objet, et une contrainte de forme sur la PSF. En outre, des méthodes sont proposées pour assurer la positivité des intensités estimées, conserver le flux de l'objet, et bien estimer le paramètre de la régularisation. Quand il s'agit d'imager des spécimens épais, la phase de la fonction pupille, due aux aberrations sphériques (SA) ne peut être ignorée. Dans la seconde partie, il est montré qu'elle dépend de la difference à l'index de réfraction entre l'objet et le milieu d'immersion de l'objectif, et de la profondeur sous la lamelle. Les paramètres d'imagerie et la distribution de l'intensité originelle de l'objet sont calculés en modifiant l'algorithme AM. Due à la nature de la lumière incohérente en microscopie à fluorescence, il est possible d'estimer la phase à partir des intensités observées en utilisant un modèle d'optique géométrique. Ceci a été mis en évidence sur des données simulées. Cette méthode pourrait être étendue pour restituer des spécimens affectés par les aberrations sphériques. Comme la PSF varie dans l'espace, un modèle de convolution par morceau est proposé, et la PSF est approchée. Ainsi, en plus de l'objet, il suffit d'estimer un seul paramétre libre.

Mots-clés: microscopie confocale à balayage laser, fonction de flou, déconvolution aveugle, approche bayésienne, maximum de vraisemblance, algorithme EM, variation totale, maximum *a posteriori*, minimisation alternatée, aberrations sphériques

Dedicated to my family and Chariji.

Acknowlegments

I would like to first thank my thesis advisor, Dr. Laure Blanc-Féraud, who accepted and hosted me as an intern, and later as her doctoral student at INRIA Sophia-Antipolis. In spite of her busy engagements, she always took the time for me, and was never hesitant to discuss any idea I had, however absurd it might seem. Her gentle guidance, friendly and informal nature provided a very good ambiance for learning. I would like to also acknowledge Dr. Josiane Zerubia, Ariana research team director, who always enlivened us with her boundless energy. Be it a research, administrative or personal problem, she had the knack of always handling them with her wisdom. I would like to take a leaf out of Josiane's book for her round-the-clock dedication to the project and to her family. It was our former project secretary Mrs. Corinne Mangin, who assisted me for more than three years with the administrative activities in INRIA and France. One could always bank on her thoughtfulness. Although she left Ariana to pursue other engagements, she will be never forgotten.

At Ariana, I am obliged to Dr. Caroline Chaux (now at Université Paris-Est, France), our team's basketball trainer and coach, for keeping her office doors and mind open for game or research strategies. After her departure from Ariana, it was my dear friend, Mr. Aymen El-Ghoul, who lent a very patient listening ear to all my frustrations and tribulations. I envy his curiosity, naïvety and his boldness towards research. He has the softness of the mind that can readily mold into any good thought. Many hear but very few listen as well as Aymen which makes any discussion with him fair and lively.

The transition from the signal processing application of power system to microscopy is never an easy road. Today, if I am able to better understand fluorescent microscopes, I owe it to my collaborators Drs. Bo Zhang (formerly with Pasteur Institute, and now with the Philips medical research, France), Zvi Kam (Weizmann Institute, Israel) and Gilbert Engler (INRA Sophia-Antipolis, France). In spite of the distance, both Zvi (Tziki) and Bo were very patient to all my email queries and gave timely reviews. In Tziki, I found a complete scientist. Thorough in his understanding and providing valuable insights as short statements. Often it required quite a lot of introspection to understand these very

significant scientific statements. He brought cheer to those around with his infectious and cherubic smile. Gilbert showed me that to find a solution to a research problem, it is often important to ask the right question. To a first timer, it is easy to misjudge him due to his humble manners and immense humor. One would never guess the true potential he hides within himself. I am also indebted to him for the many samples he prepared, experiments performed and images he obtained for testing my algorithms in this thesis. I thank also Dr. Jean-Christophe Olivo-Marin, from the Pasteur Institute, France, for ensuring the success of this Franco-Israel collaboration, and accepting to be the member of my Ph.D. jury. I am also grateful to Prof. Alain Dieterlen (Lab.El, IUT Mulhouse, France) for inviting me to his laboratory, and the courtesy showered upon me. Mulhouse might be known for its lovely Christmas market, but perhaps lesser known for its hospitality. I also take this moment to gratefully acknowledge Dr. Olivier Haeberle, Mr. Tal Kenig (formerly with Technion, Israel), and Prof. Arie Feuer (Technion, Israel) for several interesting discussions. Additionally, my sincere gratitude goes to Drs. Stéphane Noselli and Fanny Serman from the Institute of Signaling, Development Biology & Cancer UMR 6543/CNRS/UNSA for some of the images presented in this thesis. My sincere appreciation goes to all the members who agreed to be part of my jury and their reviewing my thesis. I am personally indebted to reviewers Dr. Philippe Ciuciu (LNAO, CEA, France) and the president of the jury Prof. James B. Pawley (University of Wisconsin, Madison, USA). I thank them for taking the time to critically review this thesis and providing their invaluable comments which gave further insights on this subject.

This research was funded by the P2R Franco-Israeli Collaborative Research Program. I would like to thank INRIA for supporting this Ph.D. through a CORDI fellowship. I would also like to thank my current and former colleagues at INRIA: Vikram Sharma, Csaba Benedek, Daniel Menasche, Denny Oetomo, Peter Horvath, Neismon Fahé, Avik Bhattacharya, Pierre Weiss, Alexandre Fournier, Olivier Zammit, Farzad Kamalabadi, Giuseppe Scarpa, Gabriel Ducret, Alexis Baudour, Daniele Graziani, Maria Kulikova, Ahmed Gamal Eldin, Guillaume Perrin, Dan Yu, Saloua Bouatia-Naji, Mats Eriksson and many others whom I might have unknowingly missed out. They made my stay at INRIA a very memorable one. Over the three years at INRIA, I had several office mates: Ting Peng, Rami Hagege, Stig Descamps, Nabil Hajj Chehade, and Vir-

ginie Journo. I thank them for providing a nice working ambience

Leaving the best for the last. Words would not be enough to recognize the role that my friends and family have played in my life. My parent's support in every decision I took, and their complete faith made me strive for higher goals. They were by my side encouraging me in the most disappointing situations as well as in my most successful moments. They provided the strong foundation that I could safely fall back to when needed. If you find me nice, I squarely lay the blame on my upbringing. My critics provided the inner eye for introspection. The first in the line would be my sister Radhika Pradeep, and then probably Leïla Preiss. Their observations made with love and good humor helped me work on them. Kapil Kulakkunnath and Bharathi Balasubramanian are a great couple whose goodness of heart influence everyone around them, including me. Sanjeev etta planted in me, one evening, 20 years back, probably the first seed of scientific curiosity. I will miss my associates at the Nice ashram who embraced me as one of their own, and made me appreciate the subtleties of the French language and culture. This thesis would not have been started or completed had it not been the complete support of my dear sensei Chariji. He steers me and provides for under all circumstances.

Contents

List of Tables

List of Figures

Abbreviations

3D	Three dimension
2D	Two dimension
AFP	Actual focal position
AU	Airy units
BD	Blind deconvolution
BS	Background subtraction
CCD	Charge-coupled device
CG	Conjugate gradient
CLSM	Confocal laser scanning microscope
CLT	Central limit theorem
COG	Center of gravity
DFT	Discrete Fourier transform
EM	Expectation maximization
FWHM	Full-width at half maximum
GFP	Green fluorescent protein
IBD	Iterative blind deconvolution
IDFT	Inverse discrete Fourier transform
iff	if and only if
LSI	Linear space invariant
LSM	Laser scanning microscope
LSNI	Linear space noninvariant
MAP	Maximum a posteriori
ML	Maximum likelihood
MLEM	Maximum likelihood expectation maximization
MLE	Maximum likelihood estimate
MRP	Median root prior
MSE	Mean squared error
NA	Numerical aperture
NFP	Nominal focal position
OTF	Optical transfer function
p.d.f	Probability density function

PMT	Photomultiplier tube
PSF	Point-spread function
SNR	Signal to noise ratio
SA	Spherical aberration
viz.	namely
WFM	Widefield microscope
w.r.t	with respect to

Notations

$(\cdot * \cdot)$	Linear space invariant convolution
$(\cdot)^*$	Complex conjugation operation
$\hat{(\cdot)}$	Estimate
$\|\cdot\|_p$	ℓ_p norm with $p \geq 1$
α	Semi-aperture angle of the objective lens
α_{λ_o}	Object prior's partition function parameter
a_g	Shape hyperparameter for the hyperprior
$A(\cdot)$	Apodization function
b	Background signal
b_g	Scale hyperparameter for the hyperprior
$\mathscr{B}$	Fourier-Bessel transform or Hankel transform of zero-order
β_o	Sigmoid function steepness parameter
d	Depth below the cover slip or nominal focal position
D	Diameter of the CLSM pinhole
$\delta(\cdot)$	Dirac-delta function
Δ_{xy}	Radial sampling size
Δ_z	Axial sampling size
$\mathbf{E}(\cdot,\cdot)$	Time-varying electric field
$\mathbb{E}(X)$	Expectation of a random variable X
ϵ	Algorithm convergence factor
$\epsilon_{(\cdot)}$	Permitivity of a medium
$\mathscr{F}(\cdot)$	Fourier transform
$\mathscr{F}^{-1}(\cdot)$	Inverse Fourier transform
$\mathbf{G}(\cdot)$	Gradient matrix for a function
γ	Reciprocal of the photon conversion factor
$\Gamma(a_g, b_g)$	Gamma distribution with parameters a_g and b_g
h	Point-spread function of an imaging system
$\mathbf{H}(\cdot,\cdot)$	Time-varying magnetic field
$\mathbf{H}(\cdot)$	Hessian matrix for a cost function
$i(\cdot)$	Observed image

Continued on Next Page…

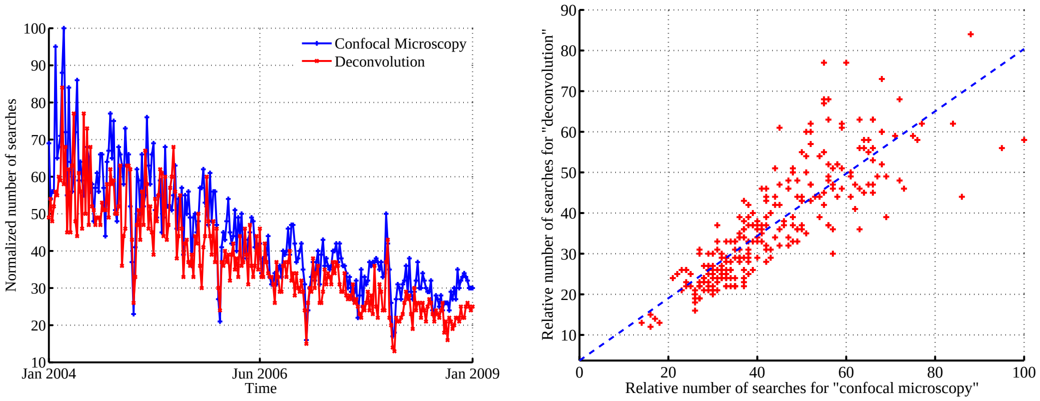
$\mathbf{I}_d$	Identity matrix
j	Imaginary unit of a complex number
J_0	Bessel function of the first kind of order zero
$\mathbf{k}$	Vector of coordinates in the frequency space
λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength
$\mathscr{J}(\cdot)$	Energy function
(μ,σ^2)	Mean and variance of a normal distribution
$\mu_{(\cdot)}$	Permeability of a medium
n	Index of iteration for the estimation algorithm $\in \mathbb{N}^+$
$n_{(\text{med})}$	Refractive index of a medium
$\mathbf{n}$	Outward normal to a surface
N_b	Number of basis components
$\nabla(\cdot)$	Gradient of a vector field
$\nabla^2(\cdot)$	Laplacian of a scalar field
$\mathscr{N}_g(\cdot)$	Additive Gaussian noise
$\mathscr{N}_p(\cdot)$	Poisson distribution
(N_x,N_y)	Number of pixels in the radial plane
(N_z)	Number of axial slices or planes
$o(\cdot)$	Specimen or object imaged
$\boldsymbol{\omega}_{(\cdot)}$	Parameter vector to be estimated
Ω_s	Discrete spatial domain
Ω_f	Discrete frequency domain
$p(\cdot)$	Probability density function
P	Pupil function
$\Pr(\cdot \mid \cdot)$	Conditional probability
$\mathscr{P}(\cdot)$	Voxel-wise Poissonian process
$\boldsymbol{\Phi}(\cdot)$	Set of basis functions
q	Average photon flux
(ρ,ϕ,z)	Cylindrical coordinates
(ρ,ϕ,θ)	Spherical coordinates
$\left(\sigma_r^2,\sigma_z^2\right)$	Lateral and axial variances for a 3D Gaussian function
σ_n^2	Gaussian noise variance

Continued on Next Page…

Σ	Diffracting aperture
$\boldsymbol{\Sigma}$	Covariance matrix
$\sum_{\mathbf{x}}(\cdot)$ or $\sum\limits_{\mathbf{x}}(\cdot)$	Discrete summation over $\mathbf{x}$
Θ	Parameter space ($\in \mathbb{R}$)
$\mathbb{V}(X)$	Variance of a random variable X
φ	Optical phase difference
φ_d	Defocus phase term
φ_a	Aberration phase term
w_l	Weights or basis coefficients
$\mathbf{x}$	Vector of coordinates in object/image space (2D or 3D) ($\in \mathbb{R}^2$ or $\in \mathbb{R}^3$)
z	Axial distance in the image space ($\in \mathbb{R}$)
Z_n^0	Zernike polynomials of zero kind and order n

Preface

Fluorescent light microscope offers the unique opportunity of observing live cells in action and permit cell biologists to study molecular and cellular mechanisms. Confocal light microscope, a type of fluorescent microscope, has especially advanced in recent years allowing the imaging of specimens at tissue and cellular levels. It is widely used in several disciplines, from cell biology and genetics to microbiology and developmental biology. It has also several clinical applications that allows rapid diagnosis and thus assist with early therapy. It is particularly useful for localizing particles, tracking cell motility, qualitative, and quantitative analysis. In comparison to conventional widefield microscope, it offers higher axial resolution, better contrast, and sensitivity to single molecule fluorescence. Recent advances in laser technology and computational power added impetus to its further development. In the confocal microscope, a diffraction-limited laser spot excites the sample, and the image is formed by collecting only the in-focus fluorescence light. A series of optical sections are combined to visualize the three dimensional (3D) structure of the sample. Although most of the out-of-focus light is rejected by the pinhole, even with a useable physical pinhole size, a third of the light collected is from the out-of-focus planes. The spatial resolution of such a microscope is limited, and even under the most suitable imaging conditions, it cannot resolve lesser than 200nm. Unfortunately, this limit in resolution prohibits the viewing of sub-cellular structures below this size. Deconvolution is a computational technique used to reverse the effects of diffraction and to remove the out-of-focus fluorescence. It is thus no surprise that the recent internet search trends shows a very high correlation between the keywords ‘confocal microscopy’ and ‘deconvolution’ (*cf.* Fig. 1). From the figure, one might be tempted to quickly dismiss the decrease in the search trend gradually over the years as the loss of interest in this subject. However, such a conclusion would be incorrect, for this search trend plot is normalized. A similar decrease was also noticed in the proportion of internet users interested in science (if a search keyword is a reflection of a user’s interest in a subject). The reason for this is that with internet becoming pervasively used, the number of internet users have exponentially increased, but the growth of the scientific community has not been proportionally



(a) (b)

Figure 1: (a) Internet search trend on the keywords 'deconvolution' and 'confocal microscopy' between the period 2004–2009 in the context of biology and medicine (b) scatter plot between the two searches showing a linear correlation (dotted line is the line of least square fit). Data source: Google Trends.

fast.

If we would like to reconstruct the true object from the observation by deconvolution, it requires the knowledge of the degradation process; the degradation being mathematically given by the imaging system's point-spread function (PSF). Given the PSF, it is then possible to computationally restore the object's true intensities by deconvolution. However, the theoretical calculation of the PSF, from acquisition parameters, is often inaccurate as these parameters might vary with conditions. An empirically estimated PSF suffers from the same difficulty as well apart from being noisy. The solution is that the restoration has to be done in a blind manner from a single observation. In Fig. 2, the number of scientific citations per year is shown for the subject of deconvolution and blind deconvolution. It is worthwhile noting that this problem has attracted a considerable academic and industrial interests recently.

Often the biologist is faced with a difficult situation of choosing the objective lens, obtaining good 3D resolution, fast image acquisition, minimal photodamage or bleaching, and minimal aberrations. Such an ideal situation is almost never met and biologists find trade-offs on these aspects of imaging. The aim of this thesis is to bring us closer in achieving these objectives by using computational methods. With the introduction of the new generation of multicore processors with high memory, and graphical processing units (GPU), the power

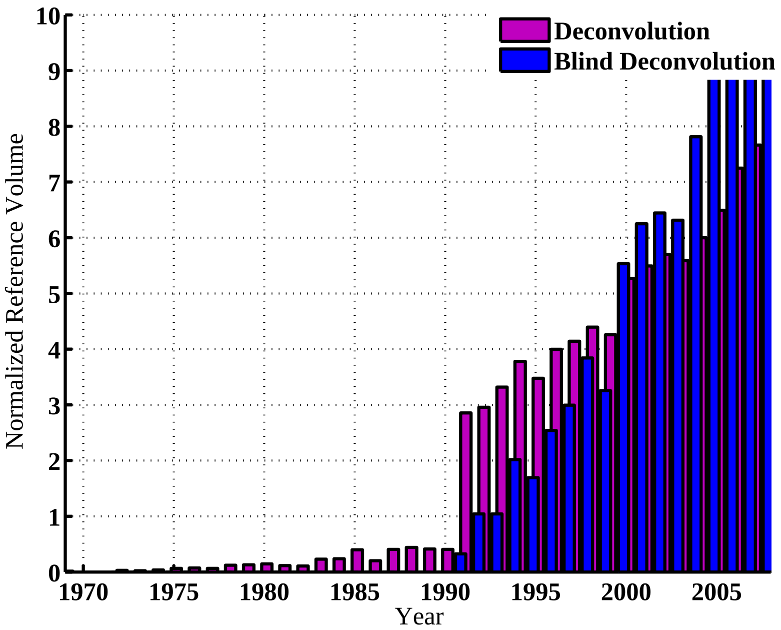


Figure 2: Annual scientific citations for the subject of deconvolution and blind deconvolution. (Data source: ISI web of knowledge).

of parallel computation is inexpensive and ubiquitous. This has provided the necessary computational stimulus to break the diffraction barrier.

In this thesis, we integrate the theoretical developments on the *direct* problem of modeling the optics of the microscope, and on the *inverse* problem of restoring the object and PSF identification. For the direct model, the physical model of the PSF's is discussed by taking into account the phenomenon of diffraction and spherical aberrations. These analytical expressions are relevant, and parsimonious because it is simplified by the insistence on the phase. The noise models in fluorescence microscopes are discussed and a clear justification is presented for not choosing the Poissonian model over the traditional linear Gaussian model. The framework adopted for the reversal is that of Bayesian inference with a correspondence between the *a posteriori* distribution and an energy functional that is minimize by an estimator. The estimation techniques proposed here are validated on simulated and real data.

The organization of this thesis and our essential contributions are as follows. In Chapter 1, we introduce fluorescent microscopes and compare the imaging properties of a widefield and confocal. We look at the fundamental difficulties faced in optical sectioning microscopy, and formalize the mathematical model for the image formation at the objective and at the detector. As stray illumination affects the observed image as background fluorescence, we provide a method to estimate it directly from the images without using the dark current images. Some quality measures are defined that are used to compare the simulated restoration results with the true object. We define the observation process as a perceptual inference problem and propose the estimation of the properties of an underlying scene from an image. We choose the Bayesian hypothesis as it provides a natural framework for modeling this inference. The algorithms for blind and non-blind deconvolution are reviewed in Chapter 2 based on this framework. A joint maximum a posteriori (JMAP) is proposed for the problem of estimating the underlying scene and the system PSF. As the blind deconvolution problem is under-determined, it is difficult to uniquely recover the PSF or the object from the observation. In order to assist the JMAP, a prior is proposed on the object and the PSF. As biological specimens are complex and rich in structure, it is important to understand them if we wish to recover them from the observation. Our belief on the object is captured by using a Gibbs' distribution, and a total variation functional is used as the energy term. The analytical PSF modeling based on Stokseth's approximation is introduced in the Chapter 3. We extended this diffraction-limited PSF model to also encompass the spherical aberrations (SA) occurring due to refractive index mismatch. To simplify the estimation of the PSF, a spatial approximation of the diffraction-limited model in terms of a three-dimensional separable Gaussian is proposed. Similarly, the PSF with SA is also approximated so that in the limiting case, the problem is reduced to estimating two or three of the free pupil function parameters. A simple numerical algorithm for the theoretical PSF model is also given which could be implemented by using discrete Fourier transforms. Chapter 4 is dedicated to the estimating the imaged object and the microscope PSF using the JMAP. A new prior on the object is introduced to that naturally handles negative intensities arising during the deconvolution. As the regularization parameter varies with observation, a new approach is proposed to learn it from the observed images. An alternate minimization (AM) algorithm is finally proposed to solve the problem of estimating two unknowns from a single cost function. This AM algorithm is tested on images of degraded phantom objects, microspheres and real data. We arrive at a conclusion interesting for biologists: our method of blind deconvolution allows for a two fold improvement in the spatial resolution. The scope of this chapter is restricted to restoring images from a CLSM given the spatial invariance nature of the diffraction-limited PSF. If empirically obtained PSFs are available, they cannot be directly used as they are often low in contrast, noisy and larger than the true PSF. Theoretically calculated PSFs are inadequate as they are not microscope specific. In Chapter 5, we modify the JMAP algorithm in order to restore the phase of the pupil function by reversing the roles of the object and the PSF. Once the parameters of the phase are estimated, the PSF could be generated numerically by using the analytical PSF model. The process of generation of the synthetic object is explicitly provided for validation, and the preliminary results show the need for a good model of the PSF. Otherwise there is the risk of providing objects less usable than the observation data. The blind parameter estimation estimated the refractive indices close to their true values, but the error on the estimation of the nominal depth of focus induces a small error in the position of the object in volume imaged. Finally, in Chapter 6, this method is extended to estimate the object and PSF under aberrations. In the presence of SA, the PSF is depth varying and the spatial invariance assumption is nullified in the axial direction. As a result, a new observation model is proposed, based on a piecewise convolution process, to overcome this difficulty. We proceed by assuming that the volume is juxtaposed by various strata in which the PSF is invariant. The AM approach is adapted so that the parameters of the PSF and the object could be estimated from the observation in the presence of SA.

CHAPTER 1

Introduction

> *"The eye of a human being is a microscope, which makes the world seem bigger than it really is."*
>
> *-Gibran Khalil Gibran (Lebanese-American artist, poet and writer)*

Light microscopes provide cell biologists the possibility of examining subcellular activities in live samples with minimum disturbance to their movements. However, the road to achieving finer spatial image resolution is riddled with problems. The physics of the light captured by the lens and the detector efficiency essentially limits the quality of the observed images in optical fluorescence microscopy. The goal of this chapter is to familiarize with these fundamental difficulties. In Section 1.1, we briefly introduce the imaging capabilities of fluorescence microscopes and in Section 1.2, we further explore the above sources of distortions. The journey towards overcoming these obstacles can be said to be partially accomplished if we are able to satisfactorily derive mathematical models for them. In Section 1.3, we formalize the problem with a physical image formation model at the microscope objective and at the emission-photon detector. As the background fluorescence can affect the performance of restoration algorithms, a method is proposed in this chapter for its estimation from observed images. Finally, we define a few quality measures in Section 1.4 that are used for validating the restoration results in the later chapters of this thesis. The limitations in the microscope optics is introduced here but a detailed discussion is left for the following chapter.

1.1 Fluorescence Microscopy

Fluorescence microscopes are optical instruments capable of obtaining three dimensional (3D) image sections of a specimen by focusing a high intensity monochromatic laser beam. The image is created by scanning the laser across the sample in a raster formation. Every point in the sample, on excitation, acts as a secondary light source either naturally or because of the expression of a specific labeling protein (such as the green fluorescence protein or GFP). The expression results from a phenomenon called *fluorescence*, where the laser excited electrons in the protein molecule, when relaxing back to their native ground states, emit light of wavelength longer than the incident beam. To achieve maximum fluorescence intensity, the fluorochrome is usually excited at the peak wavelength of the excitation curve, λ_{ex}, and the emission is selected at the peak wavelength of the emission curve, λ_{em}. The red-shift in the emitted light (Stokes shift) in comparison to the excitation light facilitates in designing filters that can separate them. Each dye has a distinct excitation and emission curve that are well separated from that of all the other dyes; consequently, only the peak emission intensity of each dye needs to be measured in order to obtain accurate quantization. By changing the objective to focus at different depths inside the specimen, and by collecting the emitted fluorescence at each plane, one can visualize the topologically complex cells, tissues and embryos in 3D.

In Fig. 1.1, we see a simple schematic of a single photon (1p) confocal laser scanning microscope (CLSM) (Minsky [1988]; Inoué [2006]), where the emission field energy is collected by placing a photomultiplier tube (PMT) at the position of the emission beam focus. The partially reflecting dichroic mirrors separate the fluorescence light that comes through the objective and the unwanted reflected excitation light. The difference from the classical fluorescent microscopes such as a widefield microscope (WFM) is that, in the CLSM, a circular pinhole is added before the detection stage. This pinhole restricts the total amount of light collected to the plane that is in focus (as shown by solid line E in the schematic in Fig. 1.1; dotted line represents the out-of-focus planes). Thus, the light rays that are incorrectly aligned with the pinhole are eliminated from the final image. Unlike in the original design of Minsky, the commercial adaptation of a CLSM uses galvanometer mirrors to tilt the laser beam as it passes through the back focal plane of the objective rather than moving the specimen stage. This prevents undue vibrations in live specimen imaging and permits one to obtain image sections without motion related aberrations.

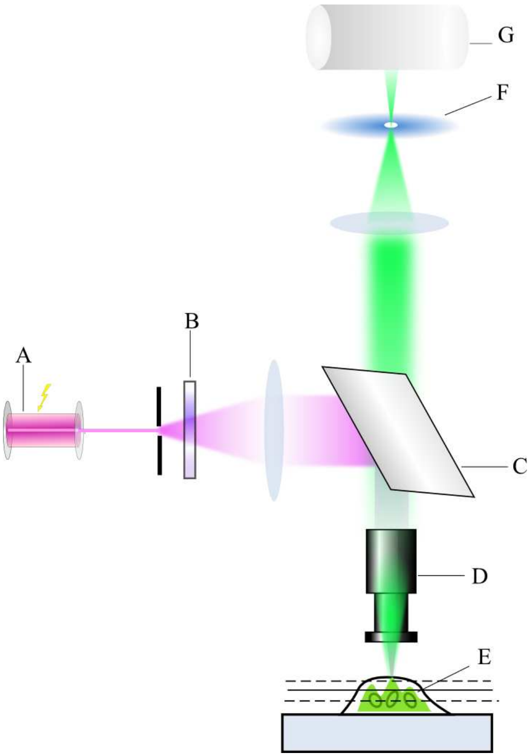


Figure 1.1: Schematic of a confocal laser scanning microscope. A) Laser, B) excitation filter, C) dichromatic mirror, D) objective lens, E) in-focus plane of the specimen, F) pinhole aperture, G) photomultiplier tube (©Ariana-INRIA/I3S).

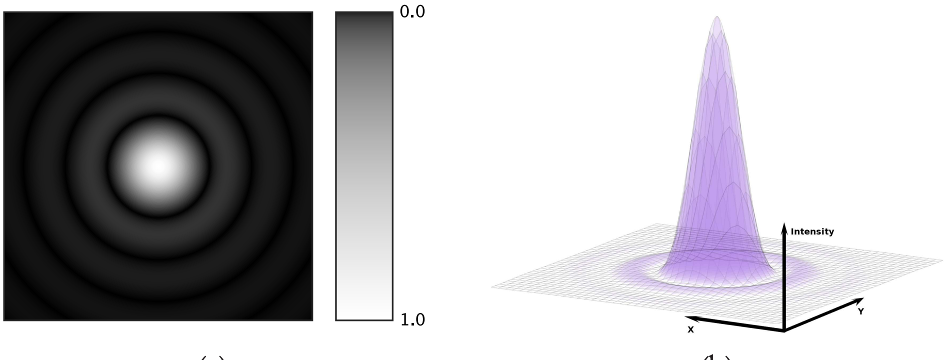
1.2 Fundamental Limits in Imaging

In this section we discuss primarily two main problems that limit the object resolving capability of a fluorescence microscope.

1.2.1 Optical Distortions

Diffraction-Limited Imaging

The optical system of a microscope is inherently diffraction limited (Pawley [2006]; Born & Wolf [1999]) and the image of a point source, the point-spread function (PSF), displays a lateral diffractive ring pattern (expanding with defocus) introduced by the finite-lens aperture. This is because when light from a point source passes through a small circular aperture, it does not produce a bright dot as an image, but rather a diffused circular disc known as Airy disc surrounded by



(a) (b)

Figure 1.2: Illustration of the diffraction pattern observed for a circular aperture as (a) intensity image scaled between $[0, 1]$ and (b) in 3D with the intensity along the z axis (ⓒAriana-INRIA/I3S).

much fainter concentric circular rings (see Fig. 1.2). This example of diffraction is of great importance because many optical instruments (including the human eye) have circular apertures. If this smearing of the image of a point source is larger than that produced by the aberrations of the system, the imaging process is said to be diffraction-limited, and that is the best resolution which can be obtained from that size of aperture. As the objective constitutes an important part of CLSM, the quality of the image, and its resolution is dependent on the quality of the lens used, its numerical aperture (NA), and the wavelength of excitation light used.

Out-of-focus Fluorescence

In addition to diffraction-limits, the optically sectioned images obtained from a uniformly illuminated 3D object are often affected by some out-of-focus fluorescence contributions. Secondary fluorescence from the sections away (as shown by dotted lines in the schematic) from the plane of focus often interferes with the contrast and resolution of those features that are in focus (as shown by the solid line in the schematic in Fig. 1.1). Let us take up the classical wide-field microscope (WFM) as a case for comparison against the CLSM. For the sake of simplicity, if we assume that their detectors are the same, then a WFM could be seen as a CLSM but with a fully open pinhole. The WFM can collect more light even from the deeper sections of a specimen but the data are sometimes rendered useless as there is a significant amount of out-of-focus blur. The maximum intensity in each plane decreases as $O(z^{-2})$ (*cf.* Zhang [2007]), with z being the axial distance from the source. A completely closed pinhole (diameter < 1 Airy units (AU)[1]) on the other hand, confines the light detected only to the in-focus plane but at the expense of imaging high-contrast, low SNR (signal dependant noise) images. For typical pinhole sizes, the maximum intensity from a point source in CLSM decreases as $O(z^{-4})$ (*cf.* Zhang [2007]) and the loss of in-focus intensity inhibits imaging of weakly fluorescent regions. We remind that even with a useable pinhole diameter of 1AU, 30% of the light collected is from the out-of-focus regions.

Aberrations

Under ideal conditions, a high NA objective lens focuses the incident planar wavefront to a spherical wavefront. However, under practical situations, the refracted wavefront so produced has to go through several optical elements and also through the specimen before it is detected. In Fig. 1.3, we show the aberrated wavefront arising at the exit pupil and the reference sphere[2]. By comparing the emerging wavefront with the reference sphere, the amount of aberrations and its severity could be established. Aberrant wavefront means that the final observed images will be distorted as well. While there are many aberrations that occur, we restrict our analysis to spherical aberration (SA) as this is the dominant and the

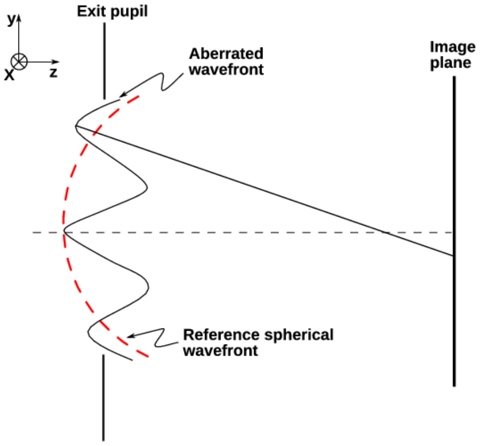


Figure 1.3: Aberrated wavefront at the exit pupil. (©Ariana-INRIA/I3S).

---

[1] $1\text{AU}=(1.22\lambda_{ex})/\text{NA}$, where NA is the numerical aperture of the objective

[2] A wavefront that has the shape of the reference sphere will come to focus at the center of this sphere.

most observable form for CLSM. In this case, the rays that are farther away from the center are refracted more than those closer to the axis. Although SA can be corrected by carefully designing the objective lens, it is only effective if certain, precise conditions are met regarding the size and refractive index (RI) of every element between the focus plane and the camera. When imaging cells in aqueous media, this is seldom the case, particularly when they are imaged with a coverslip (on an inverted microscope with no aberration correction). Even when using a water lens with correction collar for coverslip thickness, these aberrations are very common. In fact, next to the low quantum efficiency (QE) of the detector (discussed in next subsection), it is definitely the greatest practical limitation to high performance in live-cell confocal microscopy. We will discuss more about SA in the Chapters 3, 5, and 6.

Optical Resolution

We can define the lateral resolution limit for a WFM with a perfect objective and finite NA by the Rayleigh criterion (*cf.* Born & Wolf [1999]) as

$$r^{\text{wfm}}_{\text{lateral}} = 0.61 \frac{\lambda_{\text{em}}}{\text{NA}} . \tag{1.1}$$

If the distance between two closely spaced point sources is lesser than $r^{\text{wfm}}_{\text{lateral}}$, then they cannot be resolved. As $r^{\text{wfm}}_{\text{lateral}}$ is the distance between the principal intensity maximum and the first intensity minimum, two equal intensity sources are considered to be just resolved when the maxima of one coincides with the minima of the other. The spherical waves issued from the wavefront interfere not only in the image plane but also throughout the 3D space. Consequently, the image of the point source located in the object plane is a 3D diffraction pattern, centered on the conjugate image of the point source located in the image plane. The commonly used axial resolution for the confocal configuration is given below (*cf.* Sheppard [1988]):

$$r^{\text{wfm}}_{\text{axial}} = 0.885 \frac{\lambda_{\text{em}}}{n_i - \left(n_i^2 - \text{NA}^2\right)^{\frac{1}{2}}} , \tag{1.2}$$

with n_i representing the RI of the imaging medium between the coverslip and the front lens. It describes the distance between the maximum intensity of the central bright region and the first point of minimum intensity along the z-axis. If the illumination and fluorescence emission wavelengths are approximately the same, the confocal fluorescence microscope Airy disk size is the square of the widefield microscope Airy disk. Thus, the lateral (and axial) extent of the point spread function is reduced by about 30% compared to that in the WFM. Because of the narrower PSF, the separation of points required to produce acceptable contrast in the CLSM is reduced to a distance approximated by (*cf.* Dey, et al. [2004])

$$r^{\text{clsm}}_{\text{lateral}} = \frac{1}{(2)^{\frac{1}{2}}} r^{\text{wfm}}_{\text{lateral}}, \tag{1.3}$$

$$r^{\text{clsm}}_{\text{axial}} = \frac{1}{2} r^{\text{wfm}}_{\text{axial}}. \tag{1.4}$$

1.2.2 Noise Sources

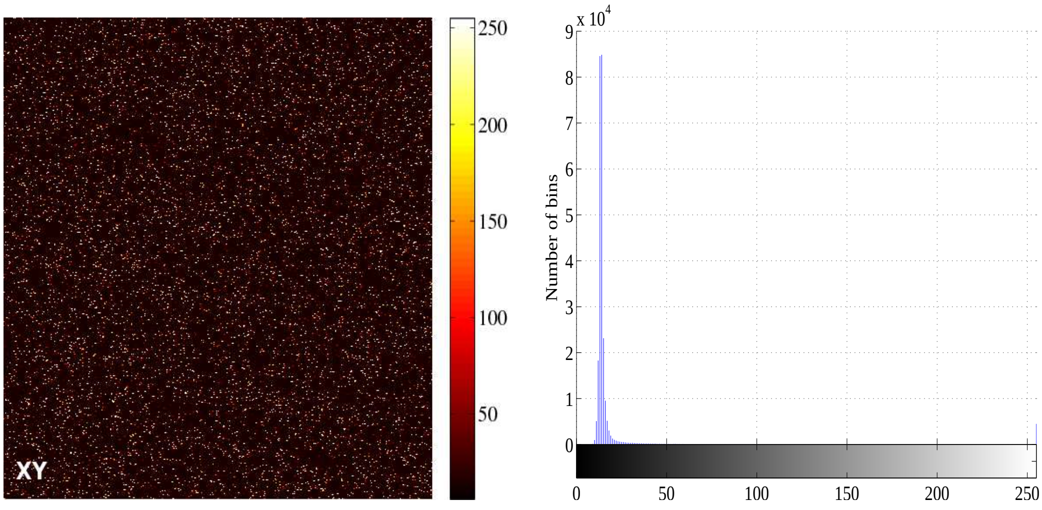
The sources of noise in a digital microscope are either the signal itself or the digital imaging system. Since we deal with CLSMs having PMT as its principal detector element, we will discuss only two kinds of noises, photon or shot noise and dark noise. Due to the quasi-random nature of the noise, knowledge of its sources and mechanics helps us to better model them. For the sources of noise in other detectors like charge-coupled device (CCD) of a WFM, we suggest the following references: Stevens, et al. [1994] and Zhang [2007].

Shot Noise

Shot noise occurs when the energy carrying photons exhibits detectable statistical fluctuations in the measurement. These fluctuations are particulary noticeable in the image when the acquisition is under low illumination conditions and the number of photons reaching the detector are small. In addition, the PMTs used for detection usually have a lower QE or detection rate in comparison to the modern day CCDs. Supposing that the average photon flux is q_s, the statistical variation in the observed photon number N_s is described as $N_s \sim \mathscr{P}(q_s)$. When the photon flux is high, N_s will be asymptotically normally distributed with both the mean and the variance equal to q_s. This noise is signal intensity dependent, which makes separation of the signal from the noise a very difficult task.

Dark Noise

Under non-ideal imaging conditions, there is another type of noise generated as a result of the dark current. Dark current occurs due to thermionic emissions in the dynodes, leakage currents, field emissions, electronic emission by cosmic rays and sometimes stray indoor illuminations. Although very small, if the detector gain is large, its contribution to the final signal is significant. This dark noise N_d has an average dark flux q_d, and $N_d \sim \mathscr{P}(q_d)$. Fig. 1.4(a) shows the signal observed in a Zeiss LSM 510 microscope when there is no excitation source, with PMT gain adjusted to get the dynamic range, and when it is operating in the "photon counting mode". The detector gain and the offset are set so that the there is no saturation or clipping of intensities. The mean and the variance of the dark current image were estimated using a maximum likelihood (ML) estimate. The mean of the background signal was about 22.3 ± 5% intensity units (IU) and variance 16.98. When the gain is minimum, the mean was found to be 5.2482 ± 5% IU and variance 7.25. In practice, the dark noise can be minimized by cooling the detector tube.



(a) (b)

Figure 1.4: (a) Measured dark current images in the absence of any excitation or sample, and the settings adjusted for permissible detector gain and amplifier offset range (ⓒINRA), and (b) histogram of the dark image. The lateral pixel size, Δ_{xy}, is 89.98μm with 8-bit gray levels.

Poissonian Assumption

From the histogram of Fig. 1.4(b), we observe that the distribution is approximately Poisson in nature. Thus, the total observed photon number, $N_p \sim \mathscr{P}(q_s + q_d)$, could be assumed to follow only a Poisson distribution (*cf.* Mandel [1979]) and by neglecting the *readout* noise. If we characterize the noise by square root of the variance, then the signal-to-noise ratio (SNR) is given by

$$\mathrm{SNR} \approx \frac{q_s + q_d}{(q_s + q_d)^{\frac{1}{2}}} = (q_s + q_d)^{\frac{1}{2}} \,. \tag{1.5}$$

1.3 Mathematical Formulation

1.3.1 Physical Model

When the total magnification of the system is not unity (say M), by the diffraction theory (discussed in Chapter 3), the observed image $i : \mathbb{R}^3 \mapsto \mathbb{R}$ (in the absence of noise) can be written in the continuous domain as

$$\begin{aligned} i(x_i, y_i, z_i) = \iiint\limits_{(x_o, y_o, z_o) \in \mathbb{R}^3} & h(x_i - M x_o, y_i - M y_o, z_i - M^2 z_o) \times \\ & o(x_o, y_o, z_o) dz_o dy_o dx_o \,, \end{aligned} \tag{1.6}$$

where $o : \mathbb{R}^3 \mapsto \mathbb{R}$ is the original specimen under investigation, $h : \mathbb{R}^3 \mapsto \mathbb{R}$ is the system PSF, and (x_i, y_i, z_i) and (x_o, y_o, z_o) are the 3D coordinates in the image and the object spaces. Normalizing the coordinates as $(x'_o, y'_o, z'_o) \equiv (M x_o, M y_o, M^2 z_o)$, we can say

$$\begin{aligned} i(x_i, y_i, z_i) = \iiint\limits_{(x_o, y_o, z_o) \in \mathbb{R}^3} & h(x_i - x'_o, y_i - y'_o, z_i - z'_o) \times \\ & o\left(\frac{x'_o}{M}, \frac{y'_o}{M}, \frac{z'_o}{M^2}\right) dx'_o dy'_o dz'_o \,. \end{aligned} \tag{1.7}$$

The above expression can be simply written as

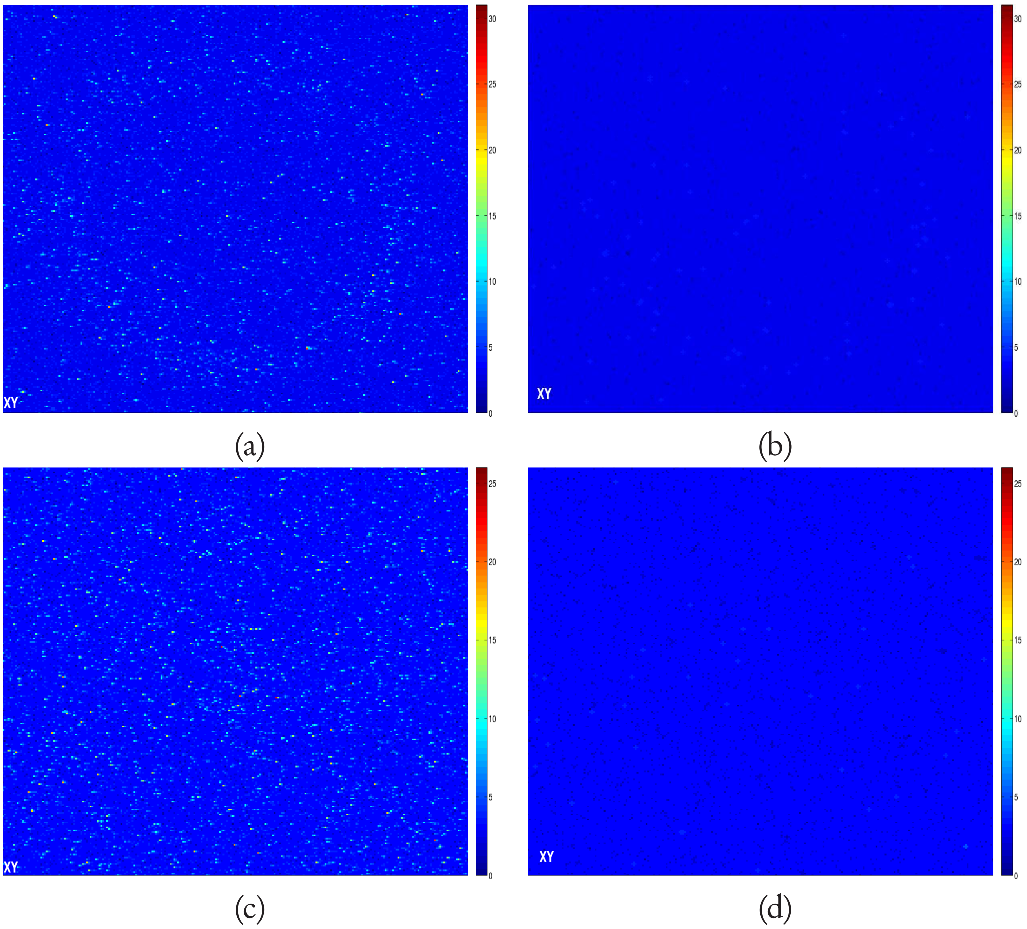
$$i(x_i, y_i, z_i) = (h * o_g)(x_o', y_o', z_o'), \tag{1.8}$$

where $o_g(x_o', y_o', z_o')$ is the geometrical optics prediction of the object and the interaction between h and o_g is a '3D convolution'. We can see that re-normalizing the object coordinates makes the model linear shift invariant (LSI) (*cf.* Goodman [2004]). From the viewpoint of computational methods, this could be inverted with the knowledge of the scanning system properties and also by information about the object being scanned. It is for this reason that the knowledge of the point-spread function (PSF), h, is of fundamental importance. The nature of the PSF for fluorescence microscope has been studied extensively in Agard [1984], Hudson, et al. [1996] and Zhang, et al. [2007]. In Chapter 3, we will introduce to the reader a theoretical model of the PSF based on the scalar diffraction theory and later its approximations adapted for blind deconvolution (BD). In the above model, we assumed isoplanatism of the microscope objective or spatial-invariance along the meridional sections, which is a valid supposition for a well aligned microscope. However, it is not often the case along the axial direction where the PSF might change with depth. We will look at this problem in more detail in Chapters 3, 5 and 6. Throughout this thesis, we will use the terms 'object' or 'specimen' interchangeably to refer to the 'biological specimen' or the underlying scene we wish to study.

1.3.2 Background Fluorescence Model

We can assume that the imaging system has been *a priori* calibrated so that there is negligible offset in the detector and that the illumination is uniform; *i.e.* no misalignment in the laser. This assumption is justified in our case as is elucidated by the following example.

In Fig. 1.5(a) and (c), we show the first and last observed optical sections obtained by imaging a fluorescent microsphere immersed in water. The first and the last slices were chosen because these are essentially sans features, and hence free from the object fluorescence. Fig. 1.5(b) and (d) shows the estimated background, $\hat{b}(\mathbf{x})$, obtained by morphologically opening the two sections by using a circular structural element (*cf.* Gonzalez & Woods [1987]). Fig. 1.6 shows the histograms for the observed image volume and the volume after the background was subtracted. When we compare the histograms, we observe how the entire histogram has been shifted towards the zero after the background subtraction (BS). What we can also infer from the background estimated section is that there is a uniformity in the illumination and no alignment problems. The mean value of the background signal from this estimation procedure was found to be about 3.13 IU. Next, the overall histogram of the image volume, in Fig. 1.6, was smoothed and a Poisson distribution was fit to the data. As before, the parameters of the distribution were estimated using a ML estimate. The empirical mean of the distribution was estimated to be lie between 3.9683 and 3.9702 IU with a 95% confi-



(a) (b)

(c) (d)

Figure 1.5: The estimated background fluorescence $\hat{b}(\mathbf{x})$ in (b) and (d) are from (a) the first slice and (c) the last slices of an observed photon counts having very sparse object fluorescence (©Ariana-INRIA/I3S, INRA). The lateral sampling for this volume is $\Delta_{xy} =$ 37.38nm and the axial sampling Δ_z is 151.45nm.

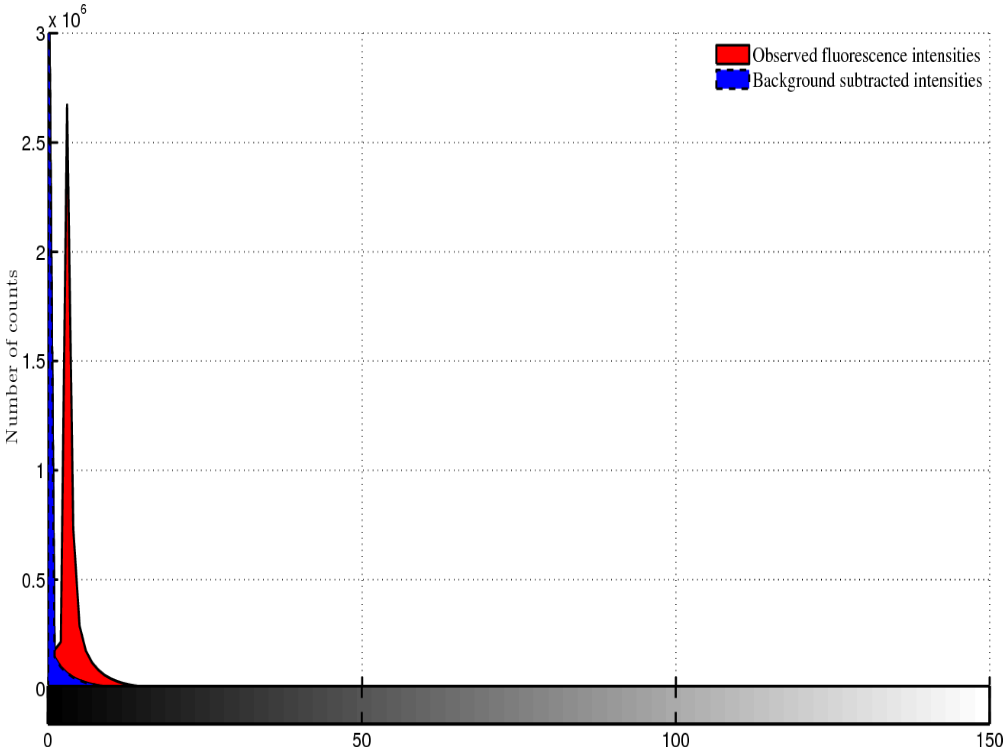


Figure 1.6: Histogram of the original and the background subtracted volume (ⓒAriana-INRIA/I3S).

dence. Since the object fluorescence was sparsely populated, we find that there is not much difference between the mean estimates by considering the overall volume or the individual section. This is valid in most of the images taken using a CLSM where the object fluorescence is sparse throughout the volume. However, we notice a significant difference between the background estimated using the dark image with full amplifier gain and the above estimation for an observed volume. This reinforces the idea that the background needs to be estimated for every observation volume, and if the object fluorescence is sparse, the estimation could be carried out on the observation. For more details on homogenous or heterogenous background estimation in fluorescence microscopy, the interested reader may refer to the following articles by van Kempen & van Vliet [2000] and Chen, et al. [2006].

1.3.3 Image Formation in an Aberration-Free Microscope

In a discrete spatial support $\Omega_s = \{(x,y,z) : 0 \leq x \leq N_x - 1, 0 \leq y \leq N_y - 1, 0 \leq z \leq N_z - 1\}$, we denote by $\mathcal{O}(\Omega_s) = \{o = (o_{xyz}) : \Omega_s \subset \mathbb{N}^3 \rightarrow \mathbb{R}\}$ the possible observable objects, and by $h:\Omega_s \mapsto \mathbb{R}$ the microscope PSF. When the imaging system has been *a priori* calibrated, there is almost negligible offset in the detector and the illumination is uniform. For an aberration-free model, the convolution is LSI, hence in a matrix notation,

$$(h*o)(\mathbf{x})=\mathbf{HO}\,, \tag{1.9}$$

and the matrix $\mathbf{H}$ is *block-circulant* (if periodic boundary conditions are used). It can be shown that if the matrix is block-circulant, it is diagonalized by the discrete Fourier transform (DFT). In addition, we will assume that the encircled or volumetric energy is constant with respect to the blurring process. This assumption is valid because blurring is a passive process and hence

$$\sum_{\mathbf{x}\in\Omega_s} o(\mathbf{x})=\sum_{\mathbf{x}\in\Omega_s}(h*o)(\mathbf{x})\,. \tag{1.10}$$

.

In the literature of restoration methods for fluorescence microscopy, there are two models used to describe the image formation process at the detector. These are the additive Gaussian white noise (AGWN) and the Poissonian model. If $\{i(\mathbf{x}):\mathbf{x}\in\Omega_s\}$ denote the observed intensity (bounded and positive) of a volume, for the Gaussian noise assumption, the observation model can be written as

$$i(\mathbf{x})=(h*o)(\mathbf{x})+w(\mathbf{x}),\,\forall\mathbf{x}\in\Omega_s\,, \tag{1.11}$$

where $w(\mathbf{x})\sim\mathcal{N}(0,\sigma_g^2)$ is an AGWN with zero mean and variance σ_g^2. If we were to approximate a Poissonian process by a Gaussian noise, the variance of the noise will depend on the mean intensity, $(h*o)(\mathbf{x})$. The high SNR case can be addressed by employing the central limit theorem (CLT) for large photon numbers, where the AGWN model performs satisfactorily. It is important to note that under low SNR, the AGWN model provides a poorer description of the fluorescence microscopy imaging than the following Poisson model.

For the Poissonian assumption, the observation model can be expressed as

$$\gamma i(\mathbf{x})=\mathcal{P}\left(\gamma([h*o](\mathbf{x})+b(\mathbf{x}))\right),\,\forall\mathbf{x}\in\Omega_s\,, \tag{1.12}$$

where $\mathscr{P}(\cdot)$ denotes a voxelwise noise function modeled as an i.i.d.Poissonian process. $b : \Omega_s \mapsto \mathbb{R}$ is a uniformly distributed intensity that models the low-frequency background signal caused by scattered photons and autofluorescence from the sample. $1/\gamma$ is the photon conversion factor, and $\gamma i(\mathbf{x})$ is the photon count at the detector.

The discussion of the mathematical approximation for image formation under aberrations is left for Chapter 6.

Nyquist Sampling

The Nyquist sampling theorem tells us that for a bandlimited system, the ideal sampling frequency should not be less than twice the cut-off frequency. For a CLSM, the radial cut-off frequency is $4\mathrm{NA}/\lambda_{\mathrm{ex}}$, and hence the radial Nyquist sampling is

$$\Delta_{xy}^{\mathrm{clsm}} = \frac{\lambda_{\mathrm{ex}}}{8\mathrm{NA}}. \tag{1.13}$$

The axial cut-off frequency is related to the depth-of-field (DOF) of the microscope (*cf.* Sheppard [1986a,b]). The DOF is the distance over which the intensity is more than half of the maximum. Since the axial cut-off frequency is $2(n_i - (n_i^2 - \mathrm{NA}^2)^{1/2})/\lambda_{\mathrm{ex}}$ the maximum axial pixel step should be

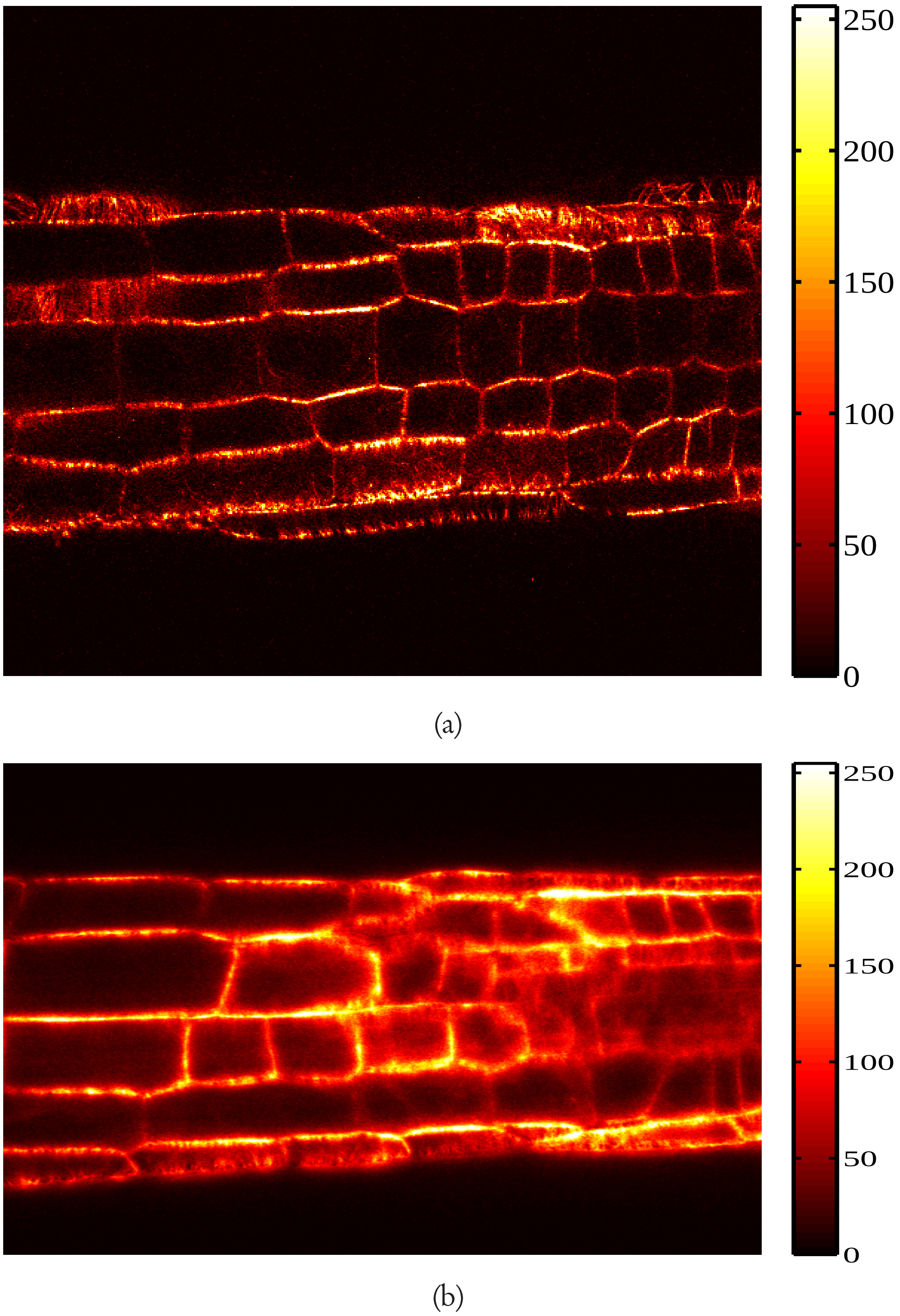
$$\Delta_{z}^{\mathrm{clsm}} = \frac{\lambda_{\mathrm{ex}}}{4(n_i - (n_i^2 - \mathrm{NA}^2)^{\frac{1}{2}})}. \tag{1.14}$$

Effect of Pinhole Size

A sample of the plant *Arabidopsis Thaliana*, immersed in water, was imaged using a Zeiss LSM 510 microscope fitted with an objective lens of 1.2NA and 63X magnification. The dimateres of the pinhole were changed, and the images were taken for four different settings: 1AU, 2AU, 5AU and 10AU. The sizes of the pinholes are the back-projected[3] values in AU. Fig. 1.7 shows the central plane of two such observed volume with minimum and maximum pinhole settings. One can also notice that with the smaller pinhole (1AU), there is more contrast and

---

[3] Back-projected diameter is the diameter of a pinhole in the object space. It is calculated as the ratio between the real physical diameter of the pinhole and the total magnification of the system



(a)

(b)

Figure 1.7: The central axial plane of an imaged *Arabidopsis Thaliana* plant in water under two different pinhole sizes (a) 1 AU and (b) 10 AU. (ⓒINRA Sophia-Antipolis)

detail in the image but the total signal collected is less as well. While, the images obtained from the larger pinhole (10AU), have larger signal (signal from the region of interest and out-of-focus signal) but loss in contrast as well due to unwanted fluorescence from the neighboring slices. There is thus a trade-off to be established between the pinhole size, amount of noise and the out-of-focus light. 1AU is the useable pinhole size for the confocal but for sizes more than 4AU, the microscope behaves like a widefield as far as the photon statistics are concerned. This is because as the total number of photons per pixel collected is relatively large, then the Gaussian model is more applicable. Fig. 1.8 shows the plot of the histogram for these two observed volumes. We remark that the histogram slowly changes from a purely Poisson distribution for the 1AU to a mixture of Poisson and Gaussian (MPG) (*cf.* Zhang [2007]) for the 10AU setting.

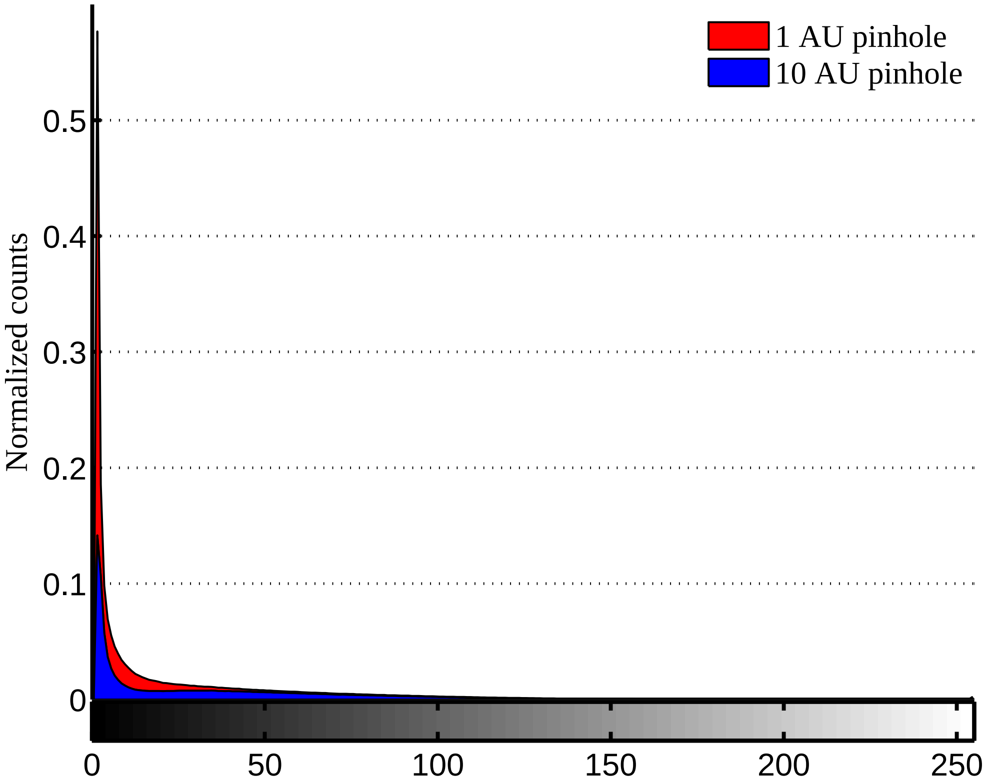


Figure 1.8: Normalized histogram for the observed specimen under two pinhole settings (©Ariana/INRIA/I3S).

1.4 Image Quality Measures

In this section, we define some measures that are used to assess the quality of the restoration algorithms proposed in this thesis. These measures could be used during simulations to quantify the results obtained by comparing the final estimate with the true object. They were proposed to mimic the human ability to distinguishing a good quality image from a poorer one. However, they come no where closer to our visual perception, and yet under the current circumstances these serve as the best judge of the quality improvement. The search for a good quality measure is still open and the authors Wang & Bovik [2009] proposed a measure based on similarity. Although this similarity index measure is introduced here, it will not be explored in this thesis.

1.4.1 Mean Squared Error

The mean squared error (MSE) is the ubiquitous preference as a metric in the field of signal processing. It gives the cumulative squared error between the restored and the original image.

Definition 1. Supposing the two images are $o(\mathbf{x})$ and $i(\mathbf{x})$, then the MSE between them is defined by

$$\mathrm{MSE}(o,i) \overset{def}{=} \frac{1}{N_x N_y N_z} \sum_{\mathbf{x}} (i(\mathbf{x}) - o(\mathbf{x}))^2 \,. \tag{1.15}$$

A more general form is based on the ℓ_p norm of the difference:

$$d_p(o,i) \overset{def}{=} \left(\sum_{\mathbf{x}} |i(\mathbf{x}) - o(\mathbf{x})|^p \right)^{\frac{1}{p}} \,. \tag{1.16}$$

The MSE however exhibits weak performance and has serious shortcomings for our application, since we deal with finding the significant perceptual difference between images. We have however used it only as a criterion to evaluate the error in the estimated PSFs.

1.4.2 Peak Signal-to-Noise Ratio

Definition 2. Peak signal-to-noise ratio or PSNR is a measure of the peak error in the images. In the literature of image processing, MSE is often converted into a PSNR measure as

$$\mathrm{PSNR} \overset{def}{=} 10\log_{10}\frac{L^2}{\mathrm{MSE}}, \tag{1.17}$$

where L is the dynamic range of allowable image pixel intensities, and MSE is as defined in Eq. (1.15).

Here, the ‘signal’ is the original object, and the ‘noise’ might either be the error in reconstruction or the error in the observation. For 8 bit images $L = 255$ while for 12 bit images $L = 4095$. The PSNR is useful if images having different dynamic ranges are being compared, but otherwise contains no new information relative to the MSE. A lower value for MSE means lesser error, and as seen from the inverse relation between the MSE and PSNR, this translates to a high value of PSNR.

1.4.3 Information-Divergence Criterion

The *information-divergence* criterion, shortly known as *i-divergence* or generalized *Kullback-Leibler* measures the discrepancy between two non-negative distributions. It was proved to be consistent (Jiang & Wang [2003]) with the axioms of Csiszar [1991].

Definition 3. Consider the two non-negative distributions o and i. $\mathrm{IDIV} : \Omega_s \times \Omega_s \mapsto \mathbb{R}$ is a function that measures the discrepancy or the distance between $i : \Omega_s \mapsto \mathbb{R}^+$ and $o : \Omega_s \mapsto \mathbb{R}^+$ as

$$\mathrm{IDIV}(o|i) \overset{def}{=} \sum_{\mathbf{x}\in\Omega_s} o(\mathbf{x})\log\left(\frac{o(\mathbf{x})}{i(\mathbf{x})}\right) - \sum_{\mathbf{x}\in\Omega_s}[o(\mathbf{x}) - i(\mathbf{x})]. \tag{1.18}$$

This is the dominant measure when it comes to evaluating deconvolution algorithms as the most popular, nonlinear multiplicative restoration algorithm (*cf.* Subsection 2.1.2), could also be derived by minimizing this criterion. However, this is not a magical solution to the problem of visual quality assessment, but they are still better than the MSE. It is important to mention here that this measure is asymmetrical and $\mathrm{IDIV}(o|i) \neq \mathrm{IDIV}(i|o)$, and the individual distributions need to be strictly positive, i.e. $o(\mathbf{x}) > 0$ and $i(\mathbf{x}) > 0$, $\forall \mathbf{x} \in \Omega_s$. The minimum $\mathrm{IDIV}(o|\hat{o})$, the better is the estimation procedure. A perfect deconvolution algorithm yields the $\mathrm{IDIV}(o|\hat{o})$ to be zero.

1.5 Conclusion

In this chapter, we introduced the optical sectioning properties of two fluorescent microscopes. The diffraction-limited nature of the microscope objective, out-of-focus blur, aberrations, and statistical noise can affect an observed image's resolution. A mathematical formulation is derived for image formation at the objective and at the detector, in the absence of lens aberrations. Often a low frequency background fluorescence, when present, affects the restoration process. This background changes with experiments, and has to estimated either from a single specimen-independent observation section or by using the complete observation volume. It was shown that when the fluorescence is sparse in the volume, both these methods yielded similar results. Some image quality measures were defined for evaluating the restoration approaches to be presented in Chapters 4, 5 and 6.

CHAPTER 2

State of the Art

> “*If you want to understand today, you have to search yesterday.*”
>
> *-Pearl S. Buck (American writer)*

The optical resolution of a confocal microscope, as defined in previous chapter, is often limited to the spot size to which the excitation beam is focused. This size is typically of the order of the wavelength of the light used. This defines the spatial limit to which the light can focus in a microscope. There are many sub-diffraction techniques that aim to break this barrier. In this chapter, we review some existing methods of non-blind and blind deconvolution, as applied to microscopy, that aim to achieve this as well. It is difficult to make an exhaustive review of all the literature available in this field. So, we chose only those that resemble the problem we are handling or methods that raised considerable interests on this subject. As the framework is based on the Bayesian hypothesis, we believe it is general enough to encompass the methods cited in literature. We also show how some of the approaches listed in this chapter could be considered as special cases of the application of Bayes’ theorem. In the end of this chapter, we list some recent advances in sub-diffraction techniques in microscopy.

2.1 Deconvolution Algorithms

As the process of imaging could be mimicked by convolving the structure of an object with a PSF, deconvolution is a mathematical procedure that aims to reverse this process to obtain the object. This reasons from the assumption that the PSF is the same as that when the image was obtained, and that the signal and the PSF data are free from noise. When analyzing the published literature on the subject of image deconvolution for 3D fluorescence microscopy, we found that primarily they differ in the initial mathematical models they choose to describe the observation, the model on the object, and their assumptions on the system's PSF. As a consequence, the optimization approach that each of them adopt for the restoration process becomes different. A survey on deconvolution algorithms was carried out earlier by Meinel [1986]; Carasso [1999]; Cannell, et al. [2006]; Sarder & Nehorai [2006] and more recently by Pankajakshan, et al. [2008b]. The most significant methods for fluorescence microscope are summarized in the Tab. 2.1. In this subsection, we will only highlight the differences between these algorithms by pointing out the advantages and drawbacks of each method. We divided the following few subsections primarily based on the observation model described in Section 1.3 and then the optimization method employed. We will assume for the time being that the PSF *h* is known (either empirically or numerically calculated).

2.1.1 Nearest and No Neighbors Method

A simplistic approach to deconvolution is the nearest neighbors (NN) by Agard [1984]. This method assumes that the most blurring within a focal plane is due to the light scattered from its two neighboring planes. They assume that the intra-plane PSF is normally much smaller than the inter-plane PSF. So, if $\hat{o}(\mathbf{x})$ is the estimated object its z^{th} slice can be estimated as

$$\hat{o}(\mathbf{x}) = \mathscr{F}_{2\text{D}}^{-1}\Big\{ (\, I(k_x,k_y,z) - \lambda_o \,[\, H(k_x,k_y,z-1) I(k_x,k_y,z-1) + \\ H(k_x,k_y,z+1) I(k_x,k_y,z+1) \,]\,) H^{-1}(k_x,k_y,z) \Big\}, \forall \mathbf{x} \in \Omega_s\,, \quad (2.1)$$

where $H(k_x,k_y,z-1)$ and $H(k_x,k_y,z+1)$ are the two dimensional (2D) optical transfer function[4] (OTF) for the two neighboring planes of the in-focus plane. $\lambda_o \in \mathbb{R}$ is an empirical measure that adjusts the contribution of the two neighboring planes to the central plane. Monck, et al. [1992] modified the above approach and proposed their no neighbors approach. The no neighbors method assumes that the PSF is axially symmetrical, and that the blurred neighbors are approximated by the blurred central plane. The influence of noise is assumed to

---

[4] The PSF and the OTF are related by the Fourier transform.

Algorithms	References
No noise	
Nearest neighbors	Agard [1984]
No neighbors	Monck, et al. [1992]
Inverse filter	Erhardt, et al. [1985]
Gaussian noise	
Solution as linear equation	
Tikhonov	Tikhonov & Arsenin [1977]
Regularized linear least squares (RLS)	Preza, et al. [1992]
Wiener filter	Tommasi, et al. [1993]
Solution as iterative algorithm	
Jansson van Cittert (JVC)	Agard [1984]
Nonlinear least squares (NLS)	Foskett & Grinstein [1990]
Poisson noise (PN)	
Maximum likelihood (ML) for PN	Holmes [1988]
Maximum *a posteriori* (MAP) for PN	Joshi & Miller [1993]; Van Kempen, et al. [1997], Verveer, et al. [1999]; Figueiredo & Nowak [2003], Dey, et al. [2006]

Table 2.1: List of deconvolution algorithms in the literature of microscopy, classified by the type of noise handled.

be negligibly small and ignored during computations. Although these two are very classical approaches, we list them here because some of the current commercial softwares still use them in their deconvolution toolbox. Since they assume contribution from only two of the central plane's neighboring slices, it is fast to compute (only $2N_z$ number of 2D forward and inverse Fourier transforms), and knowledge of the full PSF or the OTF is not necessary. The disadvantage of these methods are that, in the presence of an analytical PSF/OTF model, it only uses partial information, and hence the deconvolution could be said to be axially incomplete for the volume.

2.1.2 In a Bayesian Framework

In this subsection, we will use the Bayes' theorem to develop a framework that can encompass most deconvolution methods. From the Bayesian hypothesis, the posterior probability is

$$\Pr(o|i) = \frac{\Pr(i|o)\Pr(o)}{\Pr(i)}, \tag{2.2}$$

where $\Pr(i|o)$ is the likelihood function for the specimen and it specifies the probability of obtaining an image i from an object, and $\Pr(o)$ is a p.d.f-the prior-from which o is assumed to be generated (*cf.* Subsection 2.1.3). By using the Bayesian formula above, a rigorous statistical interpretation of regularization immediately follows. $\hat{o}$ is obtained by using the maximum *a posteriori* (MAP) estimate or by minimizing the cologarithm of the *a posteriori* as

$$\hat{o}(\mathbf{x}) = \underset{o\geq 0}{\operatorname{argmax}} \Pr(o|i) \tag{2.3}$$

$$= \underset{o\geq 0}{\operatorname{argmin}}(-\log[\Pr(o|i)]). \tag{2.4}$$

As $\Pr(i)$ does not depend on o or h, it can be considered as a normalization constant, and it shall hereafter be excluded from all the estimation procedures. The minimization of the cologarithm of $\Pr(o|i)$ in Eq. (2.4) can be rewritten as the minimization of the following joint energy functional:

$$\mathcal{J}(o|i) = \underbrace{\mathcal{J}_{\text{obs}}(i|o)}_{\text{Image energy}} + \underbrace{\mathcal{J}_{\text{reg},o}(o)}_{\text{Prior object energy}}. \tag{2.5}$$

$\mathcal{J}_{\text{obs}}: \Omega_s \mapsto \mathbb{R}$ is a measure of fidelity to the data and it corresponds to the term $\Pr(i|o)$ from the noise distribution. It has the role of pulling the solution towards the observation data. We make a decision about the underlying scene based on this cost function, and it specified as well the penalty paid by the system for producing an incorrect estimate of the scene. $\mathcal{J}_{\text{reg},o}: \Omega_s \mapsto \mathbb{R}$ corresponds to the penalty term $\Pr(o)$ on the object that ensures smoothness of the solution.

Direct Methods

The NN deconvolution algorithm for microscopy could not handle noise in the data. Since its advent, there have been numerous direct deconvolution methods *viz.* Wiener filter (Tommasi, et al. [1993]), inverse filter (Erhardt, et al. [1985]), and linear least squares (LLS) (Tikhonov & Arsenin [1977]; Preza, et al. [1992]). We present here the statistical interpretation of these methods.

If we assume that observation model follows the hypothesis as in Eq. (1.11), then the likelihood becomes

$$\Pr(i|o) = \left(\frac{1}{2\pi\sigma_n^2}\right)^{\left(\frac{N_x N_y N_z}{2}\right)} \prod_{\mathbf{x}\in\Omega_s} \exp\left(-\frac{(i(\mathbf{x}) - [h * o](\mathbf{x}))^2}{2\sigma_n^2}\right), \tag{2.6}$$

where $*$ denotes the 3D convolution operation and σ_n^2 is the variance of the Gaussian noise.

Definition 4. For $p \geq 1$ and $p \in \mathbb{R}^+$, we define the ℓ_p norm of $o(\mathbf{x})$ as

$$||o(\mathbf{x})||_p \stackrel{\text{def}}{=} \left(\sum_{\mathbf{x}\in\Omega_s} |o(\mathbf{x})|^p\right)^{\frac{1}{p}}. \tag{2.7}$$

The cologarithm of the likelihood function above becomes

$$\mathcal{J}_{\text{obs}}(i(\mathbf{x})|o(\mathbf{x})) = ||i(\mathbf{x}) - (h * o)(\mathbf{x})||_2^2,\ \mathbf{x} \in \Omega_s\,, \tag{2.8}$$

where the terms independent of i and o were dropped from Eq. (2.6). A remark on the data fidelity term $\mathcal{J}_{\text{obs}}(o(\mathbf{x}))$ in Eq. (2.8) is that it is strictly convex and quadratic. In matrix notation, Eq. (2.8) becomes

$$\mathcal{J}_{\text{obs}}(i(\mathbf{x})|o(\mathbf{x})) = ||\mathbf{I}-\mathbf{HO}||_2^2, \tag{2.9}$$

where $\mathbf{H}$ is the block-circulant PSF matrix (assuming periodic boundary conditions). If we assume a ℓ_2 norm on the prior (to be discussed in Subsection 2.1.3 in detail), the term $\mathcal{J}_{\text{reg},o}(o)$ on the object is

$$\mathcal{J}_{\text{reg},o}(o) = ||\mathbf{RO}||_2^2. \tag{2.10}$$

We can show that the gradient of Eq. (2.8), $\mathcal{J}_{\text{obs},i}(i(\mathbf{x}), o(\mathbf{x}))$, can be calculated as

$$\nabla_o \mathcal{J}_{\text{obs},i} = h(-\mathbf{x}) * h(\mathbf{x}) * o(\mathbf{x}) - h(-\mathbf{x}) * i(\mathbf{x}), \tag{2.11}$$

where $h(-\mathbf{x})$ is the Hermitian adjoint of $h(\mathbf{x})$. The minimization of Eq. (2.8) with respect to o gives the following explicit solution:

$$\hat{o}(\mathbf{x}) = f(\mathbf{x}) * i(\mathbf{x}), \forall \mathbf{x} \in \Omega_s, \tag{2.12}$$

where $f(\mathbf{x})$ is the inverse filter function. The above equation may also be written in terms of the matrix notation as

$$\hat{\mathbf{O}} = \mathbf{FI}. \tag{2.13}$$

The inverse filter matrix, $\mathbf{F}$, for the different algorithms are as defined in Tab. 2.2. ϵ_h in Tab. 2.2, is a small positive constant that bounds the higher frequencies and

Method	Inverse filter ($\mathbf{F}$)	References
Truncation	$\begin{cases} \mathbf{H}^{-1}, \forall \lvert H(\mathbf{k})\rvert \geq \epsilon_h \\ 0, \text{ otherwise} \end{cases}$	Erhardt, et al. [1985]
Wiener filter	$\mathcal{F}^{-1}\left(\frac{H^*(\mathbf{k})}{\left(\lvert H(\mathbf{k})\rvert^2 + \left(\frac{P_n(\mathbf{k})}{P_o(\mathbf{k})}\right)\right)}\right)$	Tommasi, et al. [1993]
Linear least square (LLS)	$\left((\mathbf{H}^*\mathbf{H})^{-1}\mathbf{H}^*\right)$	Gonzalez & Woods [1987]
Regularized linear least square (RLS)	$\left(\mathbf{H}^*\mathbf{H} + \lambda_o \mathbf{R}^*\mathbf{R}\right)^{-1}\mathbf{H}^*$	Preza, et al. [1992]

Table 2.2: Deconvolution by direct inverse filtering.

prevents noise amplification due to inversion. For the Wiener filter, $P_n(\mathbf{k})$ and $P_o(\mathbf{k})$ are the power spectral densities (PSD) of the noise and the object respectively. $H(\mathbf{k})$ is the 3D Fourier transform of the PSF, and $H^*(\mathbf{k})$ is its complex conjugate. Note that when $\mathbf{R} = \mathbf{I}_d$, where $\mathbf{I}_d$ is the identity matrix, the preference is for smaller norms and when $\lambda_o = 0$ the RLS reduces to the unregularized LLS solution. When $\lambda_o \neq 0$, the RLS can be thought of as a MAP estimate on observation data with AGWN. These methods are fast and have a closed form solution for the estimate $o(\mathbf{x})$. However, they intrinsically assume that the noise is AGWN and are thus valid only for high SNR images. As an exact PSF inverse does not exist for CLSM, these algorithms have difficulty in restoring the information beyond a certain cut off frequency. Another difficulty is low contrast in the estimates. In order to overcome some of these drawbacks, projection onto convex sets (POCS) was proposed as an alternative by Bregman [1965]. It was also later adopted to microscopy by Koshy, et al. [1990].

Maximum Likelihood under Gaussian Noise Assumption

Iterative image deconvolution algorithms are necessary when an exact inverse for the PSF does not exist in closed form. They are based on probabilistic techniques, and outperform direct inversion techniques, for the output has reduced noise amplification (*cf.* Pankajakshan, et al. [2008b]). They are more complex and computationally more intensive than the direct methods. However, computation is not a problem anymore with advances parallel processing and multi-core processors.

For the additive Gaussian noise model of Eq. (1.11), the maximum likelihood (ML) estimate is essentially the LLS solution for a known $h(\mathbf{x})$, with the noise being signal independent. The ML estimate (MLE) can also be thought to be obtained from Eqs. (2.4) and (2.5) when $\Pr(o) = \text{const}$. We will discuss this in more detail in Subsection 2.1.3. The object estimation is done in an iterative manner by adding a term proportional to the residual as

$$\hat{o}^{(n+1)}(\mathbf{x}) = \hat{o}^{(n)}(\mathbf{x}) + r\left(\hat{o}^{(n)}(\mathbf{x})\right)\left[i(\mathbf{x}) - (h * \hat{o}^{(n)})(\mathbf{x})\right], \forall n \geq 1, \tag{2.14}$$

where $r\left(\hat{o}^{(n)}(\mathbf{x})\right) = r_0\left[1 - 2/b_0|\hat{o}^{(n)}(\mathbf{x}) - b_0/2|\right]$ (r_0 and b_0 are constants). The algorithm is initialized with $\hat{o}^{(0)}(\mathbf{x})$ equated to the observation $i(\mathbf{x})$. This is known as the Jansson Van Cittert (JVC) (*cf.* Agard [1984]) algorithm. In this fixed-point iterative scheme, at every iteration, an error image is calculated by subtracting the estimated image from the recorded distorted image. To prevent negative intensities or very bright intensities, the error image is multiplied by a finite weight function that is defined over a positive intensity band. Finally, the weighted error is subtracted from the specimen estimate to obtain the new estimate. Foskett & Grinstein [1990], introduced the nonlinear least squares (NLS) algorithm that modified the JVC to ensure positivity in the output. The nonnegativity constraint guarantees that either the negative intensities are set to zero or the final specimen intensity is positive. This method however amplifies the high frequency noise at each iteration and thus requires a smoothing step at each iteration. Unfortunately, the smoothing operation does not work well for low SNR images. While JVC improves the resolution in the final estimated image, this method is not good for removing the noise. This algorithm converges quickly, but in the presence of noise, due to the residual term in the iterative step, it diverges fast.

Maximizing the likelihood in Eq. (2.6) or minimizing the energy in Eq. (2.8) leads to the following modified JVC algorithm by Agard [1984]:

$$\hat{o}^{(n+1)}(\mathbf{x}) = \hat{o}^{(n)}(\mathbf{x}) + r\left(\hat{o}^{(n)}(\mathbf{x})\right)\left[(h^* * i)(\mathbf{x}) - \left(h^* * h * \hat{o}^{(n)}\right)(\mathbf{x})\right], \forall n \geq 1, \quad (2.15)$$

and the h^* is the adjoint of h, and $h^*(\mathbf{x}) = h(-\mathbf{x})$. We can see that this can be seen as a steepest descent (SD) kind of algorithm with a variable step size. This method is also known as the Landweber algorithm.

Maximum Likelihood under Poissonian Noise Assumption

If we accept the Poissonian hypothesis in Eq. (1.12), then the image $i(\mathbf{x})$ can be regarded as the realization of independent Poisson processes at each voxel. Hence, the likelihood for the observation is written for this approximation as

$$\Pr(i|o) = \prod_{\mathbf{x}\in\Omega_s} \frac{([(h*o)+b](\mathbf{x}))^{i(\mathbf{x})}\exp(-[(h*o)+b](\mathbf{x}))}{i(\mathbf{x})!}, \quad (2.16)$$

where the mean of the Poisson process is $[(h*o)+b](\mathbf{x})$. The data energy function to be minimized is

$$\mathscr{J}_{\text{obs}}(o(\mathbf{x}))=\sum_{\mathbf{x}\in\Omega_s}[(h*o)+b](\mathbf{x})-i(\mathbf{x})\log[(h*o)+b](\mathbf{x})\,. \tag{2.17}$$

The function of the kind in Eq. (2.17), for the Poisson case, is called as '*self-concordant functions*' (*cf.* Boyd & Vandenberghe [2004]), and are convex although not in the strict sense. In the absence of the background, we can immediately see that minimizing Eq. (2.17) is equivalent to minimizing the i-divergence criterion in Eq. (1.18), $\text{IDIV}(i|h*o)$, between the observation and the blurred object. We included the background term in the likelihood expression of Eq. (2.16). In literature, conventionally, the background is estimated, and subtracted from the observation data. The resulting negative intensities are set to zero before deconvolution. However, Dey, et al. [2004] showed that such a procedure would cause unwanted oscillations in the restoration. When tests were done on simulated data, it was noticed that the final i-divergence, $\text{IDIV}(o|\hat{o})$, is higher for background subtracted restoration than restoration without BS. The better way to handle it is to include it in the likelihood expression.

An explicit iterative multiplicative algorithm (independently derived by Richardson [1972] and Lucy [1974]) based on maximum likelihood expectation maximization (MLEM) formalism can be obtained by minimizing the energy function in Eq. (2.17) as

$$\hat{o}_{\text{ML}}^{(n+1)}(\mathbf{x})=\hat{o}_{\text{ML}}^{(n)}(\mathbf{x})\cdot\left(\frac{i(\mathbf{x})}{\left[\left(h*\hat{o}_{\text{ML}}^{(n)}\right)+b\right](\mathbf{x})}*h(-\mathbf{x})\right)^{q_o},\forall\mathbf{x}\in\Omega_s, \tag{2.18}$$

where $q_o\in[1,\,1.5]$ controls the acceleration of the MLEM (*cf.* Sibarita [2005]). $(\cdot)$ denotes the component wise Hadamard multiplication. The division here and in the rest of this thesis is assumed to be component wise. The derivation of the MLEM from Bayes' theorem, and as derived by Richardson [1972] is summarized in Appendix B.

It can be shown that for noise-free images, this EM algorithm converges to the MLE. But, in the presence of noise, as it does not have any smoothness constraints, if unchecked, they evolve to a solution that displays many artifacts from noise amplification (for examples *cf.* Pankajakshan, et al. [2008b]). One possible remedy is terminating the iteration (manually or by using a statistical criterion) before the deterioration begins. Another solution is pre-filtering the observation data before deconvolution. The general technique used to handle the Poissonian noise is to transform the data using a variance stabilizing transform (VST) such as the Anscombe transform. This transform may be defined on the observed data $i(\mathbf{x})$ as

$$g(\mathbf{x})=2\sqrt{\left(i(\mathbf{x})+\frac{3}{8}\right)}, \forall \mathbf{x}\in\Omega_s. \tag{2.19}$$

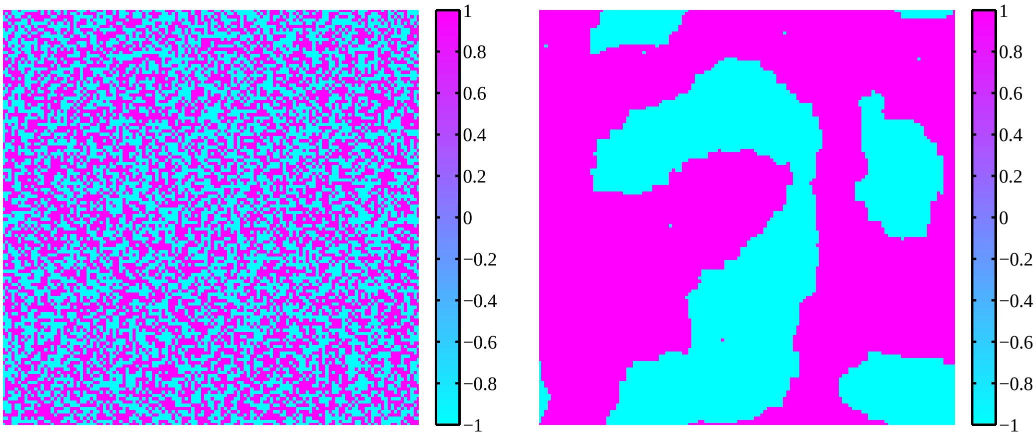
It can be shown that the noise in the resulting transformed data is asymptotically AGWN with zero mean and unit variance. The advantage of the VST is that the resulting data-fidelity term is quadratic and strictly convex. Kervrann & Trubuil [2004] used an adaptive window for denoising $g(\mathbf{x})$, while de Monvel, et al. [2001] used a Daubechies' wavelet for denoising each direction of the image stacks $i(\mathbf{x})$. There are many other denoising techniques for removing the signal-dependent noise (*cf.* Buades, et al. [2005]; Zhang [2007]; Chaux, et al. [2007]; Luisier, et al. [2010]) present in biological images. One might argue that by applying a low-pass filter as a pre-processing step before deconvolution, as by Van Kempen, et al. [1997], the results are improved in comparison to the deconvolved images with no pre-filtering. In consequence, the deconvolution algorithm is also less dependent on the object regularizing parameter, λ_o (as demonstrated in Dieterlen, et al. [2004]). However, in the light of estimating the object through deconvolution after denoising, it is difficult to analytically model the resulting denoised image as a function of the original object and the PSF. Such pre-filtering operations might influence the convolution model as it is not clear how the resulting filtered data is eventually mapped to the original object. The number of iterations for eventual convergence of the deconvolution algorithm also increases, and there is no information on the residual noise. Dupé, et al. [2008] also argued that such an approach is not recommended as the resulting transformed data does not follow the model in Eq. (1.11). We will later also see through experiments on simulated data how the estimation of the PSF, during blind deconvolution (BD), is also affected due to denoising. It is necessary to understand that such interventions are a post-hoc method of regularizing the ill-posed problem. It is a way of bringing some knowledge or smoothness to the desired solution o. We propose in the next subsection, the Bayesian interpretation of regularization problems through the MAP estimate.

Maximum *A Posteriori* under Poissonian Noise Assumption

The MAP estimate involves the simultaneous enforcing of smoothness on the object (*cf.* Section 2.1.3), and blur removal by deconvolution. Dupé, et al. [2008] used the VST to make their function convex and sparsely represent the object on a dictionary of waveforms. Their estimation involved minimizing this convex functional comprising a data-fidelity term and a non-smooth sparsity-promoting penalty over the object representation coefficients. Another approach is to directly use the MAP estimate on the observation data and modify Eq. (2.18) to also include the object regularization. We recall that the MLEM we mentioned in Eq. (2.18) can also be thought of as a specific form of the MAP estimate when the *a priori* model $\Pr(o)$ on o is uniform (or $\mathscr{J}_{\mathrm{reg},o}(o)$ is a constant). We refer the reader to the article by Demoment [1989] for more details on this subject of Bayesian interpretation of reconstruction and restoration. This second method will be our preference throughout this thesis as it lets us preserve the original data. Instead of relying on sequential algorithms to remove the noise and the blur, we use a regularization that simultaneously adds smoothness constraints in the deconvolution step. We discuss this, with more details, in Section 2.2.4. We devote the next few paragraphs on the object models and later on the survey of BD algorithms in microscopy.

2.1.3 Importance of Prior Object Model

The prior distribution $\Pr(o)$ should reflect our belief on the object o, and constraint the solution space to the most desired one. Priors express the relative probability of different scenes occurring and carries information on the scene structure. Hence, the choice of the prior assumption becomes very important in a Bayesian framework. Recently emerging research has been in the search for a good prior that models the object (*cf.* Weiss & Freeman [2007]) and the PSF. We argue that the biological specimens we investigate could be characterized using their statistics and hence they should be modeled in the stochastic sense using priors. In this subsection, we will see how incorporation of object knowledge in the reconstruction process may make the deconvolution problem well-posed.



(a) (b)

Figure 2.1: Monte Carlo simulation of Ising-Potts model with (a) $\lambda_o \approx 0$, high temperature, minimum interaction, and (b) $\lambda_o = 1$, low temperature, maximum interaction (ⓒAriana-INRIA/I3S).

Prior as Statistical Information

The ensemble model of an object class, refers to any probability distribution $\Pr(o)$ on the object space $\mathscr{O}$, and is of the following form:

$$\Pr(o) = Z_{\lambda_o}^{-1} \exp\left(-\lambda_o E(o)\right), \lambda_o \geq 0, \tag{2.20}$$

where $E(o)$ is an energy function and $1/\lambda_o$ is the Gibbs parameter. This method specifies probabilistically the relationship and interaction between neighboring voxels. We associate with each site $(x,y,z) \in \Omega_s$ of the object, a unique neighborhood $\eta_{xyz} \subseteq \Omega_s \setminus (x,y,z)$, and we denote the collection of all neighbors $\eta = \{\eta_{xyz} | (x,y,z) \in \Omega_s\}$ as the neighborhood system. If we assume that the random field, $(O = o)$, on a support Ω_s is Markovian with respect to the neighborhood system η, then, $\Pr(o_{xyz} | o_{\Omega_s \setminus x,y,z}) = \Pr(o_{xyz} | o_{\eta_{xyz}})$. o is a Markov random field (MRF) on (Ω_s, η), iff o denotes a Gibbs ensemble on Ω_s and the energy is a superposition of potentials associated to the cliques (a set of connected pixels). Hence, $E(o) = \sum_{C \in \mathscr{C}} V_C(o)$.

Fig. 2.1(a) and (b) shows the results of a Monte Carlo simulation of two scenario of 2D lattice interactions. There are several forms of the energy function, $E(o)$, based on either ℓ_1 (Dey, et al. [2006]; Pankajakshan, et al. [2007, 2009b]), ℓ_2 (Conchello & McNally [1996]), ℓ_1-ℓ_2 (Hom, et al. [2007]) norms, or entropy (Verveer, et al. [1999]), wavelets (Figueiredo & Nowak [2003]), sparsity

(Fergus, et al. [2006]) and median root priors (MRP) (Alenius & Ruotsalainen [1997]). These are tabulated in Tab. 2.3 as well.

ℓ_1 Regularization

We consider once again the images in Fig. 1.7 of Subsection 1.3.3. The gradients of these observed image volumes were numerically calculated and their histograms plotted as in Fig. 2.2. We mentioned that as the size of the pinhole is increased, the observed images have less detail due to neighboring slice fluorescence contributions. We can observe this phenomenon in the histograms of the numerical gradients, as they tend to have longer tails when the pinhole sizes are smaller in comparison to larger pinhole sizes. In summary, the objects that are in-focus tend to have longer tails in the gradients rather than objects that are out-of-focus. We will see how this information can be captured while modeling the object prior distribution in terms of the MRF.

We first define the following first-order, homogeneous, isotropic MRF, over a six member neighborhood $\eta_{\mathbf{x}} \in \eta$ (*cf.* Fig. 2.3) of the site $\mathbf{x} \in \Omega_s$,

$$\Pr[O=o] = Z_{\lambda_o}^{-1} \exp\left(-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|\right), \text{ with } \lambda_o \geq 0, \tag{2.21}$$

where $|\nabla o(\mathbf{x})|$ is a *potential functional* and λ_o is the regularization parameter. The estimation of this regularization parameter is dealt with in Section 4.3.1. From a numerical perspective, $|\nabla o(\mathbf{x})|$ is not differentiable in zero. An approach to circumvent this problem is to regularize it, and instead to consider the smooth, discrete definition as

$$|\nabla o(x,y,z)|_\epsilon = \Big((o(x+1,y,z)-o(x,y,z))^2 + (o(x,y+1,z)-o(x,y,z))^2 + (o(x,y,z+1)-o(x,y,z))^2 + \epsilon^2 \Big)^{\frac{1}{2}}, \tag{2.22}$$

where ϵ is an arbitrarily small value ($< 10^{-3}$). For the *partition function*, $Z_{\lambda_o} = \sum_{o \in \mathscr{O}(\Omega_s)} exp(-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|_\epsilon)$, to be finite, we restrict the possible values of $o(\mathbf{x})$ so that the numerical gradient $|\nabla o(\mathbf{x})|$ is also bounded. When this model is used as a prior model for the object, we have the following smoothed total

Regularization	$E(o)$	References
ℓ_2		
Gaussian	$\frac{\|\|o(\mathbf{x})\|\|_2^2}{2\sigma_o^2}$	Conchello & McNally [1996]
Tikhonov-Miller	$\sum_{\mathbf{x}} \|\nabla o(\mathbf{x})\|^2$	Tikhonov & Arsenin [1977]; Miller [1970]
Good's Prior	$\sum_{\mathbf{x}} \frac{\|\nabla o(\mathbf{x})\|^2}{o(\mathbf{x})}$	Verveer, et al. [1999]
Median Root Prior	$\sum_{\mathbf{x}} \frac{\|o(\mathbf{x})-f(\mathbf{x})\|^2}{f(\mathbf{x})}$ $f(\mathbf{x})$ is the median of $o(\mathbf{x})$	Alenius & Ruotsalainen [1997]
ℓ_1	$\sum_{\mathbf{x}} \|\nabla o(\mathbf{x})\|$	Rudin, et al. [1992]; Charbonnier, et al. [1997] Persson, et al. [2001]; Dey, et al. [2006] Pankajakshan, et al. [2008b]
$\ell_1 - \ell_2$	$\delta^2 \sum_{\mathbf{x}} \phi_o(\frac{\nabla o(\mathbf{x})}{\delta})$	Charbonnier, et al. [1993]; Chenegros, et al. [2006] Hom, et al. [2007]
Entropy	$o(\mathbf{x}) - m(\mathbf{x}) - o(\mathbf{x})\log(\frac{o(\mathbf{x})}{m(\mathbf{x})})$	Verveer & Jovin [1997]
ℓ_p norm of wavelet coefficients	$\sum_{\mathbf{x}} \|\|W(o(\mathbf{x}))\|\|_p$	Figueiredo & Nowak [2003]; Pantin, et al. [2007] Dupé, et al. [2008]; Carlavan, et al. [2009]

Table 2.3: Energy functions for a random field ($O = o$).

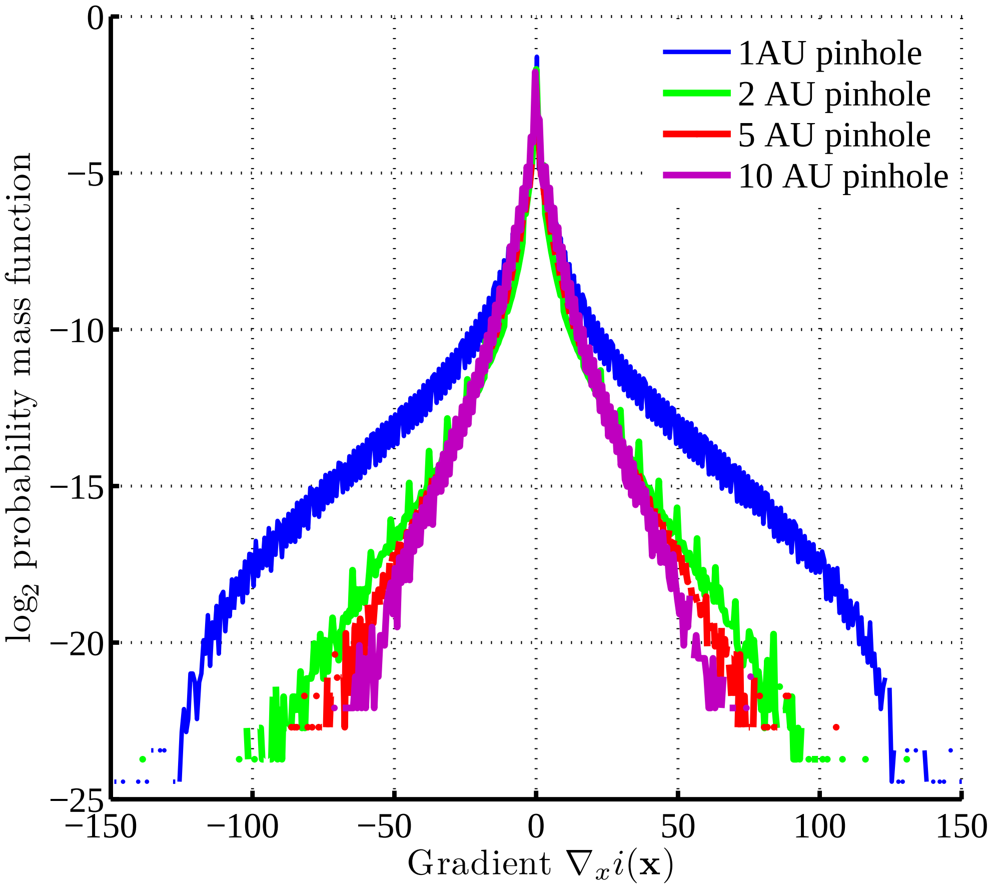


Figure 2.2: Histograms of the numerical gradients for four different confocal pinhole settings (ⓒAriana-INRIA/I3S).

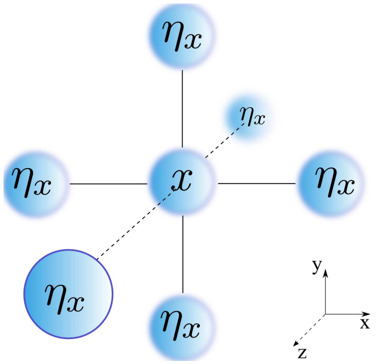


Figure 2.3: Markov random field over a six member neighborhood $\eta_{\mathbf{x}}$ for a voxel site $\mathbf{x} \in \Omega_s$ (ⓒAriana-INRIA/I3S).

variational (TV) regularization functional (*cf.* Rudin, et al. [1992]):

$$\mathscr{J}_{\text{reg},o}(o(\mathbf{x})) = \lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|_\epsilon \,. \tag{2.23}$$

Although a spectrum of regularization functionals exist in literature, we use the ℓ_1 norm because it is known for its edge preserving qualities (*cf.* Rudin, et al. [1992]; Charbonnier, et al. [1997]; Aubert & Kornprobst [2001]) and convexity. This regularization provides non-isotropic smoothing in the contour edge direction and not in the orthogonal direction of the contour. For numerical calculations, we will use a smoothed isotropic approximation of $|\nabla o(\mathbf{x})|_\epsilon$ as defined in Dey, et al. [2006], and it will be henceforth be referred simply as $|\nabla o(\mathbf{x})|$. A direct 3D extension by Charbonnier, et al. [1997] of this iterative noise removal algorithm is described by Persson, et al. [2001] for their work on 3D tomographic images. Rodríguez & Wohlberg [2009] gave a brief survey on minimization of TV type functionals. However, we use the 3D numerical implementation of the regularization algorithm for CLSM as described by Dey, et al. [2004], and in Pankajakshan, et al. [2008b]. This is summarized in Appendix D as well. The resulting functional $\mathscr{J}(o,h|i)$ is a sum of Eqs. (2.17) and (2.23). They are a combination of two functions which are self-concordant (*cf.* Boyd & Vandenberghe [2004]) and max-type (*cf.* Nesterov [2004]). Weiss [2008] showed that the above functional could be minimized with a rate of convergence of the order $O(1/\epsilon\log(\epsilon))$.

A link between the TV and the undecimated Haar wavelet soft threshold function is mentioned in Pantin, et al. [2007]. In the 1D case, these two are equivalent. However, in higher dimensions, this relation does not hold any more, but they do share some similarities. The TV introduces translation- and rotation-invariance, while the undecimated Haar presents translation- and scale-invariance.

Comments on Other Priors

A very instructive study on regularization, especially applied to fluorescence microscopy, was carried out by Verveer, et al. [1999]. Both Gaussian and Poisson models were considered for the noise in combination with Tikhonov, entropy, Good's roughness regularization and also with no regularization (MLE with flat prior). The Gaussian prior is widely used in fluorescence imaging because it prevents noise amplifications and has excellent convergence properties. Tikhonov & Arsenin [1977]; Miller [1970] introduced a regularization based on the ℓ_2 norm of the gradient of the image. However, ℓ_2 norms are isotropic as they smooth along both the tangent and the normal to the contour surface. For BD applications, a ℓ_2 kind of regularization on the object does not make the separation of the PSF and the object easy as the solution space is very similar. It has also been observed that natural image statistics rarely follow a ℓ_2 or Gaussian priors. In addition, from the example given in Fig. 2.2, it was shown that the gradients of biological specimens have heavy tailed distributions rather than Gaussian.

The MRP uses the output of a median filter as a parameter of the Gibbs prior as given in Tab. 2.3. It has the advantage that it can control noise without edge blurring. The desired object is assumed to be locally monotonic and the prior uses a simple quadratic function of the difference between the value at the central pixel and the median of its neighborhood. However, in this case, the estimate of the object is forced to be smooth, and hence can lead to loss of low-contrast information during the restoration.

Wavelet decomposition can perform quite well by analyzing the different scales separately in the deconvolution process. In literature, there are many papers that are based on deconvolution of 2D images by using wavelets. Notable among them are by Starck & Bijaoui [1994]; Figueiredo & Nowak [2001]; Jalobeanu, et al. [2003]; Willett, et al. [2003]; Vonesch & Unser [2007]. However, development of a 3D deconvolution scheme using wavelets is difficult because only an approximate of the inverse kernel is known for 3D microscopy. On the other hand, wavelet denoising can substitute as a form of regularization before or after 3D deconvolution. In this way, denoising in the wavelet scales, could be combined with regular deconvolution algorithms as mentioned in Subsection 2.1, enabling higher resolution (*cf.* Chaux, et al. [2007]).

2.2 Blind Deconvolution Algorithms

The aim of the BD for fluorescence microscopy is to answer the following simple but not so apparent question: "How does one estimate an object's true intensity and the PSF of the microscope, given only a single observed volume?"

If we forget the effect of noise and consider the observation model $(h * o)$ in the Fourier space as $\mathscr{F}(i) = \mathscr{F}(h) \cdot \mathscr{F}(o)$, several solutions answer to this problem. For example, if (h, o) is a solution, then the trivial case is that h is a Dirac function and $o = i$ or vice versa. If h is not irreducible, there exists h_1 and h_2 such that

$h = h_1 * h_2$, and the couples $(h_1 * h_2,\ o)$ and $(h_1,\ h_2 * o)$ are also solutions. Another ambiguity is in the scaling factor. If $(h,\ o)$ is a solution, then $(\tau h, o/\tau)$, $\forall \tau > 0$ are solutions too. This last ambiguity can be removed, for example, by imposing a forced normalization on h. Thus, broadly speaking, a way of reducing the space of possible solutions and to regularize the problem is to introduce constraints on h and o.

We reconsider the joint posterior probability density relationship once again using the Bayes' theorem

$$\Pr(o,h|i) = \frac{\Pr(i|o,h)\Pr(o)\Pr(h)}{\Pr(i)}, \tag{2.24}$$

where $\Pr(h)$ is the prior probability on the PSF. Conventional algorithms estimate the object and the PSF directly from the above equation by MAP,

$$(\hat{o},\hat{h}) = \underset{(o,h)\geq 0}{\operatorname{argmax}} \{\Pr(o,h|i)\}. \tag{2.25}$$

In order to simultaneously recover the two functions o and h, the following *alternate minimization* (AM) or the *zigzag* approach is usually adopted:

$$\begin{aligned} \hat{o}^{(n)}_{\text{MAP}} &= \underset{o\geq 0}{\operatorname{argmax}} \left\{ \Pr(o,\hat{h}^{(n)}|i) \right\} \\ &= \underset{o\geq 0}{\operatorname{argmin}} \left\{ -\log\left(\Pr\left(o,\hat{h}^{(n)}|i \right) \right) \right\}, \end{aligned} \tag{2.26}$$

and followed by

$$\begin{aligned} \hat{h}^{(n+1)}_{\text{MAP}} &= \underset{h\geq 0}{\operatorname{argmax}} \left\{ \Pr\left(\hat{o}^{(n)},h|i \right) \right\} \\ &= \underset{h\geq 0}{\operatorname{argmin}} \left\{ -\log\left(\Pr\left(\hat{o}^{(n)},h|i \right) \right) \right\}. \end{aligned} \tag{2.27}$$

The AM is often employed when the objective function to be minimized has two or more unknowns. In Bezdek, et al. [1987], the authors had shown that this kind of coordinate descent approach is locally convergent to a minimizer, and that the rate of convergence in each vector variable is linear. For BD, one starts with an initial guess for $\hat{h}^{(0)}(\mathbf{x})$ and obtains the alternating sequence of conditional minimizers as

$$\hat{h}^{(0)} \rightarrow \hat{o}^{(0)} \rightarrow \hat{h}^{(1)} \rightarrow \hat{o}^{(1)} \rightarrow \ldots \tag{2.28}$$

By using the Markovian property for conditional probabilities, we can equivalently say (*cf.* Chan & Shen [2005])

$$\begin{aligned} \Pr(\hat{h}^{(n+1)}|\hat{o}^{(n)}, \hat{h}^{(n)}, \hat{o}^{(n-1)}, \hat{h}^{(n-1)}, \ldots) &= \Pr(\hat{h}^{(n+1)}|\hat{o}^{(n)}), & (2.29) \\ \Pr(o^{(n+1)}|\hat{h}^{(n)}, \hat{o}^{(n-1)}, \hat{h}^{(n-1)}, \hat{o}^{(n-2)}, \ldots) &= \Pr(\hat{o}^{(n)}|\hat{h}^{(n)}). & (2.30) \end{aligned}$$

We will see more on the application of this AM algorithm in Section 4.2.

2.2.1 Marginalization Approach

We can also approach the BD problem by marginalizing o (*cf.* Jalobeanu, et al. [2007]; Blanc-Féraud, et al. [2009]) and estimate h or its vector $\theta \in \Theta \subseteq \mathbb{R}^d$. Once an estimate of the PSF or its parameters is obtained, we come back to the non-blind complete case where the PSF is known, and use the methods suggested in Subsection 2.1.2. For example, marginalizing the likelihood $\Pr(i|h)$ over all possible object $o \in \mathcal{O}$ gives

$$\begin{aligned} \Pr(i|h) &= \int\limits_{\mathcal{O}} \Pr(i, o|h)\, do, \\ &= \int\limits_{\mathcal{O}} \Pr(i|o, h) \Pr(o)\, do. \end{aligned} \tag{2.31}$$

In the Bayesian approach, all the nuisance parameters could be integrated out of the problem by using the marginalization. Hence, if the object or the PSF are parameterized, these parameters could also be integrated out of the problem. In this same manner, a marginalization could also be done over the *a posteriori* rather than the likelihood as

$$\begin{aligned} \Pr(h|i) &= \Pr(i|h) \Pr(h), \\ &= \int\limits_{\mathcal{O}} \Pr(i|o, h) \Pr(o) \Pr(h)\, do. \end{aligned} \tag{2.32}$$

However, the greatest challenge with marginalizing on o is that the probability density $\Pr(i|o,h)\Pr(o)$ in Eq. (2.32) cannot be computed in closed form. Levin, et al. [2009] proposed some alternate approximations that aim to overcome this problem but for a sparse prior.

2.2.2 *A Priori* Point-Spread Function Identification

The PSF may also be identified separately from the original image, and later used in combination with one of the image restoration algorithm that restores the object. If a confocal image is collected by using the backscattered light signal, in parallel with the fluorescent signal, the observation is filled with images of small refractile point-like objects (*cf.* Pawley [2006]). These sources in the image may be manually identified and treated (clipped and filtered) to extract the PSF (*cf.* de Monvel, et al. [2001]). This is not BD as it involves manual intervention for PSF identification from the observation images. The drawback of this kind of method is that sometimes, the point sources are not easily identified, and there is the noise factor to be handled as well.

2.2.3 Joint Maximum Likelihood Approach

For BD, we reconsider the likelihood in Eq. (2.16) now as

$$\Pr(i|o,h) = \prod_{\mathbf{x}\in\Omega_s} \frac{([(h*o)+b](\mathbf{x}))^{i(\mathbf{x})}\exp(-[(h*o)+b](\mathbf{x}))}{i(\mathbf{x})!}, \qquad (2.33)$$

where the mean of the Poisson process is given by $[(h*o)+b](\mathbf{x})$. An iterative multiplicative algorithm for the object restoration was derived in Eq. (2.18). We notice that in Eq. (2.33), the object o and h can be interchanged causing no change to the original equation. Hence, the PSF could be restored by an iterative algorithm similar to Eq. (2.18). The BD algorithm involves alternating between the object minimization and the PSF minimization as

$$\hat{o}_{\mathrm{ML}}^{(n+1)}(\mathbf{x}) = \hat{o}_{\mathrm{ML}}^{(n)}(\mathbf{x}) \left(\frac{i(\mathbf{x})}{\left[\left(\hat{o}_{\mathrm{ML}}^{(n)} * \hat{h}_{\mathrm{ML}}^{(n)}\right)+b\right](\mathbf{x})} * \hat{h}_{\mathrm{ML}}^{(n)}(-\mathbf{x}) \right)^{q_o}, \forall \mathbf{x}\in\Omega_s, \qquad (2.34)$$

followed by

$$\hat{h}_{\text{ML}}^{(n+1)}(\mathbf{x}) = \hat{h}_{\text{ML}}^{(n)}(\mathbf{x}) \left(\frac{i(\mathbf{x})}{\left[\left(\hat{o}_{\text{ML}}^{(n+1)} * \hat{h}_{\text{ML}}^{(n)} \right) + b \right](\mathbf{x})} * \hat{o}_{\text{ML}}^{(n+1)}(-\mathbf{x}) \right)^{q_h}, \forall \mathbf{x} \in \Omega_s, \quad (2.35)$$

where n is the index of iteration of the algorithm. The parameters $q_o, q_h \in [1, 1.5]$ control the convergence of the two evolutions. When q_o and q_h are unity, then we arrive at the naïve MLEM BD algorithm (*cf.* Holmes [1992] with band-limiting constraints on the PSF), and when they are more than 1, they converge (and sometimes diverge) faster. This form of grouped coordinate descent is commonly used in BD. Most iterative BD algorithms are based on the Bayesian inference and hence draw their inspiration from the above basic form of grouped co-ordinate descent. The only distinction being the different constraints they impose on the PSF and the object.

Biggs & Andrews [1997] suggested an approach for accelerating this algorithm. Since the object estimate converges faster than the PSF estimates, acceleration was achieved by performing several iterations for the PSF after each object estimation. The number of cycles of PSF iterations to apply after each iteration of the object estimate was experimentally chosen. Given an observation data, the estimation of the parameters (q_o, q_h) or the number of iterations in Biggs' method remains still an open problem. Although this algorithm works well, it suffers from some major drawbacks. Firstly, it requires information about the experimental settings to place constraints on the spatial and spectral extent of the PSF. In many cases this might not be available or might change with experimental conditions. Secondly, in its original form, it cannot be applied when there are aberrations (for fluorescent microscope *cf.* Section 3.4) in the imaging process.

2.2.4 Joint Maximum *A Posteriori* Approach

In the classical MLEM, the number of iterations, or a threshold on the MSE$(\hat{h}^{(n-1)}, \hat{h}^{(n)})$ and MSE$(\hat{o}^{(n-1)}, \hat{o}^{(n)})$ serve as stopping criteria. This is a constraint introduced in the algorithm to prevent the algorithm from continuing until divergence. The earliest BD approach that was applied to incoherent quantum-limited imaging and application to fluorescent microscopes was by

Holmes [1992]. At each iteration of the MLEM for the PSF, a unit summation constraint is enforced. In addition, the energy of the PSF is constrained to lie within an hourglass shaped region by using a Gerchberg & Saxton [1972] type algorithm. This prevents portion of the background fluorescence $b(\mathbf{x})$ from being erroneously considered as having originated from the PSF. Finally, bandlimiting and positivity criterions are introduced. The earliest such approach was proposed by Ayers & Dainty [1988] and is referred to as iterative blind deconvolution (IBD) (*cf.* Alg. 1).

For the PSF estimation, in confocal microscopy, most algorithms either assume that the PSF has a uniform distribution with certain constraints like bandlimits and positivity or follow a certain distribution. Many methods use an iterative approach to estimate the PSF and the object with no prior information on the object (Holmes [1992]; Michailovich & Adam [2007]). Markham & Conchello [1999] worked on a parametric form for the PSF and developed an estimation method utilizing this model. The difficulty in using this model for our application is that the number of free parameters to estimate is large and the algorithm is computationally expensive. Hom, et al. [2007] proposed a myopic deconvolution algorithm that alternates between iteration to deconvolve the object and estimate the PSF. In order to myopically reconstruct the PSF, they introduce a constraint on the OTF. Whatever be the prior knowledge, the important point that should be noted for application to BD is that the prior terms should enable separation of the object and the PSF solution spaces. For a review of other existing methodologies in the field of BD, the interested reader may refer to the articles by Campisi & Egiazarian [2007]; Bishop, et al. [2007]; Levin, et al. [2009].

2.2.5 Non-Bayesian Approaches

A popular BD method is the *nonnegativity and support constraints recursive image filtering algorithm* or NAS-RIF (*cf.* Kundur & Hatzinakos [1996, 1998]). Although this algorithm is known to have good convergence properties, it cannot be used for deconvolving fluorescent microscope images as the PSF does not have an exact inverse. Among the non-iterative BD algoritms, we found APEX in Carasso [2001, 2003] to be fastest in terms of computational time and applicability to a certain class of PSFs (low-exponent, Lévy stable probability density functions) with Gaussian noise. Although the defocus blurs fall under the above

Algorithm 1: Schema for the Iterative Blind Deconvolution Algorithm.

```
begin
    Input: Observed volume $i(\mathbf{x})\ \forall \mathbf{x} \in \Omega_s$.
    Data: Energy of the additive noise parameter $\epsilon_h$, and maximum
          number of iterations $n_{\max}$.
    Output: Deconvolved volume $\hat{o}(\mathbf{x})$, and estimated PSF $\hat{h}(\mathbf{x})$.
    Initialization: $n \leftarrow 0$, $\hat{o}^{(n)}(\mathbf{x}) \leftarrow i(\mathbf{x})$, $\hat{h}^{(n)}(\mathbf{x}) \leftarrow 1(\mathbf{x})$.
    while $n \leq n_{max}$, do
        Calculate $\hat{O}^{(n)}(\mathbf{k})$, the DFT of $\hat{o}^{(n)}(\mathbf{x})$.
        Fourier space PSF projection: $\hat{H}^{(n)}(\mathbf{k}) = \frac{I(\mathbf{k})\left(\hat{O}^{(n)}(\mathbf{k})\right)^*}{|\hat{O}^{(n)}(\mathbf{k})|^2 + \frac{\epsilon_h}{|\hat{H}^{(n-1)}(\mathbf{k})|^2}}$.
        Calculate $\hat{h}^{(n)}(\mathbf{k})$ from $\hat{H}^{(n)}(\mathbf{k})$ by IDFT.
        Real space PSF projection: Impose positivity and finite support
        constraints s. t. $\sum_{\mathbf{x} \in \Omega_s} \hat{h}^{(n)}(\mathbf{x}) = 1$.
        Fourier space object projection: $\hat{O}^{(n+1)}(\mathbf{k}) = \frac{I(\mathbf{k})\left(\hat{H}^{(n)}(\mathbf{k})\right)^*}{|\hat{H}^{(n)}(\mathbf{k})|^2 + \frac{\epsilon_h}{|\hat{O}^{(n)}(\mathbf{k})|^2}}$.
        Calculate $\hat{o}^{(n+1)}(\mathbf{x})$ by IDFT of $\hat{O}^{(n+1)}(\mathbf{k})$.
        Real Space Object Projection: Impose positivity and flux
        preservation constraints on $\hat{o}^{(n+1)}(\mathbf{x})$ s. t.
        $\sum_{\mathbf{x} \in \Omega_s} \hat{o}^{(n)}(\mathbf{x}) = \sum_{\mathbf{x} \in \Omega_s} i(\mathbf{x})$.
        Assign: $n \leftarrow (n+1)$.
    end
end
```

category, the method is not suitable for fluorescent microscopes, as the PSFs in this case are dependent on imaging conditions and the specimen.

2.3 Sub-Diffraction Microscopy

The resolution barrier posed by diffraction is not a true limit because this limit is applicable only when operating in the far-field region (with the exception of stimulated emission depletion microscopy, or STED microscopy). The near-field scanning optical microscope (NSOM) uses a light source and detector which in itself is nanometer in scale. Thus, the resolution in this case is not limited by diffraction, but by the size of the aperture used for scanning and its distance from the fluorophore. The apertureless NSOM or ANSOM overcomes the problem of limited photon counts posed by NSOM by creating some locally bright spots. The stimulated emission depletion (STED) microscopy, on the other hand, uses two overlapped laser pulses, the second of which has a zero intensity at its origin, and takes advantage of the non-linear de-excitation of fluorescent dyes. By sending an excitation pulse followed by a depletion pulse, the size of the excitation region is reduced. This is because the depletion beam is focused not as a spot but as a ring. All the excited electrons return to the ground state except the ones very close to the axis of the depletion beam. The photo-activated localization microscopy (PALM) and stochastic optical resolution microscopy (STORM) are based on the principle of photactivation of a very small selective fraction of the fluorophores present. A Gaussian model is then used to fit the PSFs produced by imaging the individual photoactivated molecules. The image is built up molecule-by-molecule, each of them being localized at different times. In WFM, structured illumination (SI) is used to improve the resolution. A grid of strips with predefined widths are projected onto the focal plane of the objective and then shifted laterally to the sample. The images obtained as a result of this lateral shifting is superimposed in the Fourier space, and the inverse Fourier transform yields the super-resolved reconstructed image in the real space. A recent development is the 4Pi microscope which is a CLSM that obtains high resolution by focusing the beam ideally on all sides and scan the object point-by-point. In Vicidomini, et al. [2009], an AM algorithm was proposed for 4Pi-microscopy. This method is involves alternatively estimating both the relative phase between the focusing beams, and the object. However, an assumption of constant phase throughout the observation, makes the model LSI.

2.4 Conclusion

The focus of this chapter is to review the available literature on deconvolution and blind deconvolution methods in microscopy. These methods could also be seen from a statistical angle by using the Bayes' theorem. Direct and iterative methods were studied as minimization of an energy function involving a data fidelity term and a prior term. To avoid amplification of the noise, and to make the problem of deconvolution well posed, a prior model on the object is necessary.

When maximizing the posterior p.d.f, the prior could be treated as a nuisance parameter and integrated out of the problem. An alternative, is to model it by using the Gibbs' distribution. An ℓ_1 norm on the regularization is suggested for the object prior due to its superior edge-preserving property. For restorations under aberration-free conditions, a parametric model of the PSF is suggested and its closed-form expression will be discussed in the following chapter. A JMAP approach is proposed and an alternate minimization procedure is derived for object and PSF estimation.

CHAPTER 3

Modeling the Microscope's Optical System

> *"Focusing isn't just an optical activity, it is also a mental one."*
>
> *-Bridget Louise Riley (English Painter)*

As a point-spread function (PSF) reflects the fluorescence activity at every point in the specimen, it constitutes the essential building block for recreating an object's intensities. The algorithms for deconvolution in Chapter 4 rely on an accurate knowledge of the PSF. This PSF can be determined either theoretically using a mathematical model or empirically by acquiring 3D image of a fluorescent microsphere (Hiraoka, et al. [1990]; McNally, et al. [1994]; Shaevitz & Fletcher [2007]). As theoretical models are easier to calculate and use, most deconvolution algorithms in the literature use them. We derive in this chapter, the Debye's scalar diffraction model for a lens system which eventually serves as the basis for obtaining an analytical expression for the PSF. In Section 3.1, we review the background theory, and in Section 3.2 formulate the diffraction integrals for a lens system. Based on these integrals, in Section 3.3, a simple method is provided to derive the diffraction-limited PSF from the pupil function. A spatial approximation of this PSF using an anisotropic separable 3D Gaussian function is discussed in Subsection 3.3.4. Spherical aberrations (SA) are the primary form of aberrations affecting fluorescence microscope images. These are discussed in Section 3.4, and a new but simple analytical expression for the PSF is given which includes SA. Although electromagnetic fields are vectorial in nature, in this thesis, we consider only the scalar properties of light. Our reasons for ignoring the vectorial nature of light is discussed at the end of this chapter in Subsection 3.5.1.

3.1 Background

In this section, we recall some fundamental definitions and results that serve as the base for the diffraction integral equations discussed later.

Maxwell and Helmholtz's Equations

If $\mathbf{E}(\mathbf{x},t)$ and $\mathbf{H}(\mathbf{x},t)$ are orthogonally time-varying electric and magnetic fields (*cf.* Fig. 3.1), in the absence of free charge, the *Maxwell's equations* are (*cf.* Born & Wolf [1999])

$$\nabla\times\mathbf{E}=-\mu_m\frac{\partial}{\partial t}\mathbf{H}, \tag{3.1}$$

$$\nabla\times\mathbf{H}=\epsilon_m\frac{\partial}{\partial t}\mathbf{E}, \tag{3.2}$$

$$\nabla\cdot\epsilon_m\mathbf{E}=0, \tag{3.3}$$

$$\nabla\cdot\mu_m\mathbf{H}=0, \tag{3.4}$$

where μ_m is the permeability, and ϵ_m is the permittivity of the medium. $\nabla\times\mathbf{E}$ gives the curl of $\mathbf{E}$, and $\nabla\cdot\epsilon_m\mathbf{E}$ gives the divergence of $\mathbf{E}$. If the medium is homogeneous, ϵ_m is constant throughout the region of propagation. The medium is said to be non-dispersive to light if ϵ_m is independent of the wavelength λ of the light used. All media are nonmagnetic hence the permeability of the medium is the same as that in vacuum.

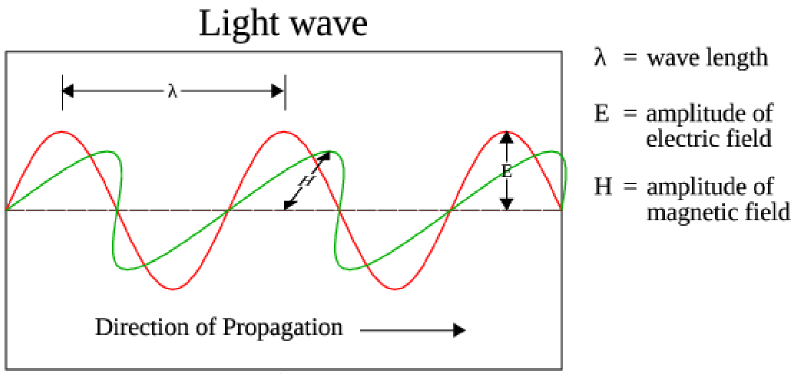


Figure 3.1: An illustration of coherent, monochromatic light as an electromagnetic wave. The electric and magnetic fields oscillate at directions orthogonal to the propagation of the light (©Ariana-INRIA/I3S).

The vector wave equation for any time varying field $\mathbf{V}(t)$ is

$$\nabla^2 \mathbf{V}(t) - \frac{n^2}{c^2}\frac{\partial^2}{\partial t^2}\mathbf{V}(t) = 0, \tag{3.5}$$

where ∇^2 is the Laplacian operator with respect to the spatial coordinates, $n_m = (\epsilon_m/\epsilon_0)^{1/2}$ is the RI of the medium, and $c = 1/(\mu_0\epsilon_0)^{1/2}$ the speed of light in vacuum. However, for non homogeneous medium, the coupling between the electric and magnetic fields cannot be rejected and hence the above equation has to be modified to include also the variation in the RI. At the boundaries, coupling is introduced between the electric and magnetic fields and in their individual components as well. The error is small provided that the boundary conditions affect an area that is only a small part of the area through which a wave may be passing.

We define a strictly monochromatic time-harmonic scalar field by

$$U(\mathbf{x}, t) = U_{\mathbf{x}}(\mathbf{x}) \exp(-j\omega t). \tag{3.6}$$

The above scalar field also satisfies Eq. (3.5), and hence the *Helmholtz equation*

$$(\nabla^2 + k^2)U = 0, \tag{3.7}$$

where $k = (2\pi n \mu_0)/c = 2\pi n/\lambda$ is the wave number, and λ is the wavelength of the light in free space ($\lambda = c/\mu_0$). If this field has no evanescent components, it can be written as a weighted sum of plane waves of the form

$$U(\mathbf{x}) = \int_{\mathbf{k}} A(\mathbf{k}) \exp(j\mathbf{k}\cdot\mathbf{x}) d\Omega, \tag{3.8}$$

where $\mathbf{k}$ is a unit vector that describes the direction of propagation of the plane waves (*cf.* Fig. 3.1).

3.1.1 Divergence Theorem

The divergence theorem or the Gauss' theorem is the higher dimensional form of the fundamental theorem of calculus.

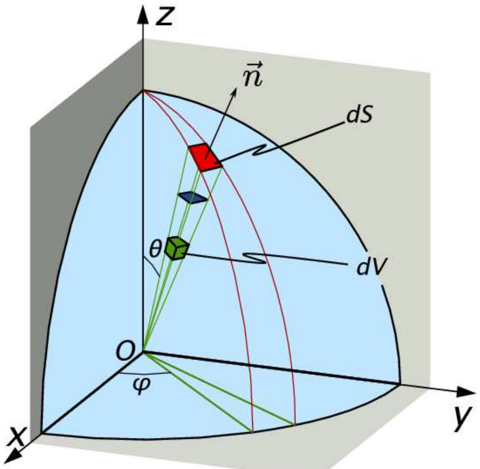


Figure 3.2: Surface and volume element depiction for Green's identity and the divergence theorem (ⓒAriana-INRIA/I3S).

Theorem 3.1.1. *Let V be a simple solid region, and S the boundary surface of V, with the positive outward orientation* **n** *(as shown in Fig. 3.2). Let* **V** *be a vector field whose component functions have continuous partial derivatives on an open region that contains V. Then*

$$\int_V (\nabla\cdot\mathbf{V})dV = \oint_S \mathbf{V}\cdot\mathbf{n}dS\,. \tag{3.9}$$

This theorem, relates the flux of a vector field through a surface S to its behavior inside the surface. For example, if the vector field acts tangential to the surface, the right hand side of Eq. (3.9) is zero. This implies that there are no flux lines flowing into the volume. This is evident.

3.1.2 Green's Second Identity

Green's identities are a set of vector derivative/integral identities that can be useful in deriving the Fresnel-Kirchoff diffraction equations and the Debye integral approximation. Since we are only interested in the second identity, it is stated as below.

Definition 5. If G and U are continuously differentiable scalar fields on $V\subset\mathbb{R}^3$, then

$$\int_V (G\nabla^2 U - U\nabla^2 G)dV = \oint_S (G\frac{\partial}{\partial\mathbf{n}}U - U\frac{\partial}{\partial\mathbf{n}}G)\cdot dS\,. \tag{3.10}$$

$G(\mathscr{P})$ is a scalar field as a function of position $\mathscr{P}$, and $\partial/\partial\mathbf{n}$ is the partial derivative along the outward normal direction in the surface element dS. It is straightforward to show that, for the scalar fields G and U, satisfying the Helmholtz equation Eq. (3.7), the left hand side of Green's second identity Eq. (3.10) is zero (also known as Helmholtz's integral theorem).

3.2 Diffraction in a Lens System

We introduce below the diffraction theory for light propagation in a homogeneous medium. In Fig. 3.3, if $\mathscr{P}_0$ is the point of observation, for diffraction by an aperture Σ, the Kirchoff, G, at an arbitrary point $\mathscr{P}_1$ that is a solution to Eq. (3.10) is

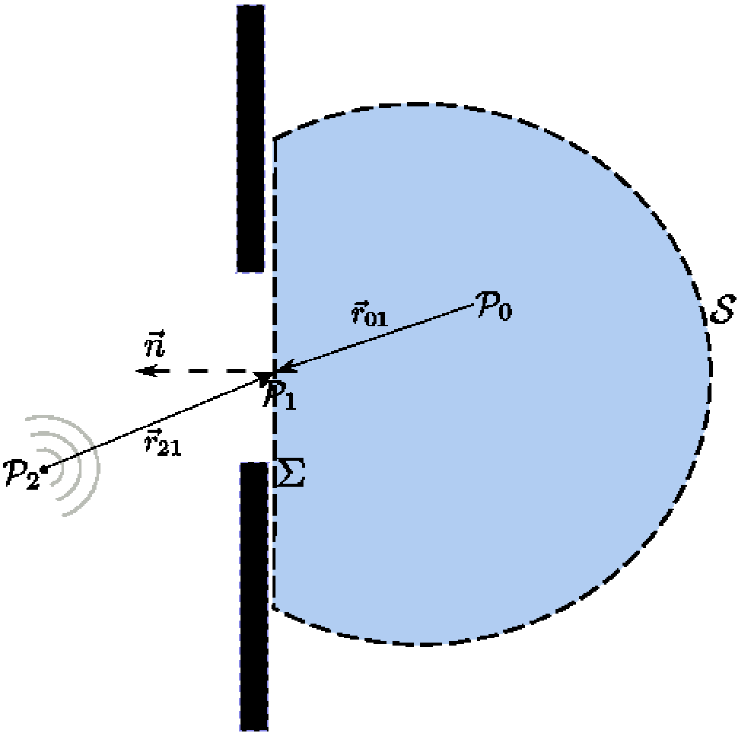


Figure 3.3: Diffraction by a planar screen illuminated by a single spherical wave (©Ariana-INRIA/I3S).

$$G(\mathscr{P}_1) = \frac{1}{r_{01}} \exp(jk_0 r_{01}). \tag{3.11}$$

where $\mathbf{r}_{01}$ is the distance from the aperture Σ to the observation point $\mathscr{P}_0$, $k_0 = 2\pi/\lambda$ is the wave number in vacuum, and $r_{01} = ||\mathbf{r}_{01}||_2$.

$$\frac{\partial}{\partial \mathbf{n}} G(\mathscr{P}_1) = \cos(\mathbf{n}, \mathbf{r}_{01}) \left(j k_0 - \frac{1}{r_{01}} \right) \frac{\exp(j k_0 r_{01})}{r_{01}}, \tag{3.12}$$

where $\mathbf{n}$ is the outward normal to Σ, and $\cos(\mathbf{n}, \mathbf{r}_{01})$ is the cosine of the angle between the normal $\mathbf{n}$ and $\mathbf{r}_{01}$.

Theorem 3.2.1. *The theorems of Helmholtz Eq.* (3.7) *and the Kirchoff Eq.* (3.11) *give the field at any point $\mathscr{P}_0$ expressed in terms of the boundary values of the wave on any closed surface $\mathscr{S}$ surrounding that point. Accordingly,*

$$U(\mathscr{P}_0) = \frac{1}{4\pi} \oint_{\mathscr{S}} \frac{\exp(j k_0 r_{01})}{r_{01}} \frac{\partial}{\partial \mathbf{n}} U - U \frac{\partial}{\partial \mathbf{n}} \frac{\exp(j k_0 r_{01})}{r_{01}} d\mathscr{S}. \tag{3.13}$$

This is the integral theorem of Helmholtz-Kirchoff and it represents the fundamental equation governing the scalar diffraction theory.

Remark 1. By the reciprocity theorem of Helmholtz (*cf.* Born & Wolf [1999]), the effect of placing the point source at $\mathscr{P}_2$ and observing at $\mathscr{P}_0$ is equivalent to placing the point-source at $\mathscr{P}_0$ and observing at $\mathscr{P}_2$.

Theorem 3.2.2. *If we assume that an aperture Σ is illuminated by a single spherical wave originating from $\mathscr{P}_2$ and that $r_{01} \gg \lambda$, $r_{21} \gg \lambda$, the disturbance at $\mathscr{P}_0$ with Kirchoff's boundary conditions is*

$$U(\mathscr{P}_0) = \frac{A}{j\lambda} \int_{\Sigma} \frac{\exp(j k_0 (r_{21} + r_{01}))}{r_{21} r_{01}} \left(\frac{\cos(\mathbf{n}, \mathbf{r}_{01}) - \cos(\mathbf{n}, \mathbf{r}_{21})}{2} \right) d\mathscr{S}. \tag{3.14}$$

Proof. By Th. 3.2.1 and applying the Kirchoff's boundary conditions we get

$$U(\mathscr{P}_0) = \frac{1}{4\pi} \int_{\Sigma} \left(G \frac{\partial}{\partial \mathbf{n}} U - U \frac{\partial}{\partial \mathbf{n}} G \right) d\mathscr{S}. \tag{3.15}$$

The fringing effects can be neglected if the dimension of the aperture is much larger than the wavelength λ of the light used. If $k_0 \gg 1/r_{01}$, then Eq. (3.12) can be written as

$$\frac{\partial}{\partial \mathbf{n}} G(\mathscr{P}_1) \approx j k_0 \cos(\mathbf{n}, \mathbf{r}_{01}) \frac{\exp(j k_0 r_{01})}{r_{01}} . \tag{3.16}$$

Substituting Eq. (3.16) in Eq. (3.15), we get

$$U(\mathscr{P}_0) = \frac{1}{4\pi} \int_{\Sigma} \frac{\exp(j k_0 r_{01})}{r_{01}} \left(\frac{\partial}{\partial \mathbf{n}} U - j k_0 U \cos(\mathbf{n}, \mathbf{r}_{01}) \right) d\mathscr{S} . \tag{3.17}$$

If we assume that the aperture is illuminated only by a single spherical wave arising from $\mathscr{P}_2$, then

$$U(\mathscr{P}_1) = A \frac{\exp(j k_0 r_{21})}{r_{21}} , \tag{3.18}$$

where $r_{21} = ||\mathbf{r}_{21}||_2$. Hence,

$$U(\mathscr{P}_0) = \frac{A}{j\lambda} \int_{\Sigma} \frac{\exp(j k_0 (r_{21} + r_{01}))}{r_{21} r_{01}} \left(\frac{\cos(\mathbf{n}, \mathbf{r}_{01}) - \cos(\mathbf{n}, \mathbf{r}_{21})}{2} \right) d\mathscr{S} . \tag{3.19}$$

□

Remark 2. The Fresnel-Kirchoff diffraction formula essentially confirms the Huygens principle. The field at $\mathscr{P}_0$ arises from an infinite number of fictitious secondary point sources located within the aperture. The secondary sources here contain amplitude and phases that are related to the illumination wavefront, and the angles of illumination and observation.

Remark 3. The Fresnel-Kirchoff diffraction approximation is similar to the Rayleigh-Sommerfield theory (*cf.* Born & Wolf [1999]) for small diffraction angles.

3.3 Diffraction-Limited Point-Spread Function Model

In incoherent imaging, the distribution of intensity in the image plane is found by integrating the intensity distributions in the diffraction images associated with each point in the object. Thus, if $o(\mathbf{x}')$ is the intensity at $\mathbf{x}'$ in the object plane, the intensity at the point $\mathbf{x}$ in the image is obtained as (ignoring the influence of noise)

$$i(\mathbf{x}) = \int\limits_{\mathbf{x}' \in \mathbb{R}^3} o(\mathbf{x}')h(\mathbf{x}-\mathbf{x}')d\mathbf{x}' \,. \tag{3.20}$$

As was mentioned before, to accurately reconstruct an object's intensity distribution o, knowledge of the PSF plays a very important role. Hence, considerable effort has been directed at characterizing their properties (*cf.* Gibson & Lanni [1989]). We summarize the most relevant models here as it helps us in later studying them for the different imaging settings.

3.3.1 Debye Approximation

From the Fresnel-Kirchoff formulation in Eq. (3.14), the Debye integral approximation for a circular aperture can be obtained as

$$U(\mathscr{P}_0) = -\frac{j}{\lambda}\int\limits_{\Sigma} \frac{\exp(jk_0(r_{21}+r_{01}))}{r_{21}r_{01}}\left(\frac{1-\cos(\mathbf{n},\mathbf{r}_{21})}{2}\right) dS \,. \tag{3.21}$$

In this we assume that the angle between the normal vector $\mathbf{n}$ and $\mathbf{r}_{01}$ is small, then $\cos(\mathbf{n},\mathbf{r}_{01}) \approx 1$. By changing the coordinates $\mathscr{P}_0 = (\rho\cos\psi, \rho\sin\psi, z)$ and $\mathscr{P}_2 = (r_{21}\sin\theta_i\cos\phi, r_{21}\sin\theta_i\sin\phi, -r_{21}\cos\theta_i)$, it is straightforward to show that (Zhang [2007])

$$\begin{aligned} U(\mathscr{P}_0) = -\frac{j}{\lambda}\int\limits_0^{2\pi}\int\limits_0^{\alpha} & A(\theta_i)\exp(-jk_i\rho\sin\theta_i\cos(\phi-\psi))\times \\ & \exp(jk_i z\cos\theta_i)\sin\theta_i \, d\theta_i d\phi \,, \end{aligned} \tag{3.22}$$

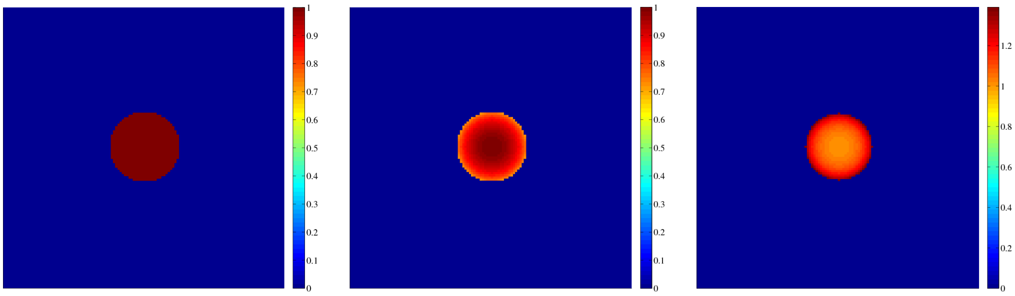
where $0 \leq \phi \leq 2\pi$, $0 \leq \psi \leq 2\pi$. The semi-aperture angle of the objective is $0 \leq \theta_i \leq \alpha$. A is the *apodization function*, $k_i = 2\pi n_i/\lambda$ is the wave number of an illumination wave, and λ/n_i is the wavelength in the medium of RI n_i. For a lens with a uniform aperture, the apodization function is radially symmetrical with respect to the optic axis and can be represented by $A(\theta_i)$. The intensity projected from an isotropically illuminating point source such as a fluorophore, on a (flat) pupil plane is bound by the energy conservation constraint. Therefore the amplitude as a function of θ_i in the pupil plane should vary as $(\cos\theta_i)^{-1/2}$ and the energy as $(\cos\theta_i)^{-1}$ (Hanser, et al. [2003]). Thus, $A(\theta_i)$ for illumination is given by

$$A(\theta_i) = (\cos\theta_i)^{\frac{1}{2}}; 0 \leq \theta_i \leq \alpha\,, \tag{3.23}$$

and for the detection as

$$A(\theta_i) = (\cos\theta_i)^{-\frac{1}{2}}; 0 \leq \theta_i \leq \alpha\,. \tag{3.24}$$

The ideal pupil functions along with the excitation and illumination pupils are shown in Fig. 3.4. Eqs. (3.23) and (3.24) are known as the "*sine condition*" and they guarantee that a small region of the object plane in the neighborhood of the optic axis is imaged sharply by a family of rays which can have any angular divergence. This constitutes an aplanatic imaging system and it exhibits 2D transverse shift invariance.



(a) (b) (c)

Figure 3.4: The amplitude of the pupil function of a CLSM for the (a) ideal, (b) illumination, and (c) emission cases (ⓒAriana-INRIA/I3S).

Remark 4. The Debye approximation holds good only if

(a) $r_{21} \gg a$, (a is the radius of the aperture),

(b) the spherical wavelets from the aperture Σ are approximated by plane wavelets,

(c) $\cos(\mathbf{n}, \mathbf{r}_{01}) \approx -1$,

(d) the NA is very large and the Fresnel number is high.

3.3.2 Stokseth Approximation

Stokseth [1969] also arrived at the approximation in Eq. (3.22) by extending the work of Hopkins [1955]. Hopkins essentially worked out an approximation for small amounts of defocus, while Stokseth's is an approximation for small and large defocus.

The PSF is defined here as the irradiance distribution in the image plane of a point source in the object plane. Diffraction, as we know it, is a far-field effect, and the light from an aperture is the Fourier transform of the aperture in the far-field. As the OTF for incoherent illumination and the PSF are related by Fourier transforms, we say

$$h(\mathbf{x}) = \mathscr{F}^{-1}(\mathrm{OTF}(\mathbf{k}))\,, \tag{3.25}$$

where $\mathbf{x} \in \Omega_s$ and $\mathbf{k} \in \Omega_f$ are the 3D coordinates in the image and the Fourier space respectively. Ω_f is the support in the Fourier space. If we consider a converging spherical wave in the object space from the objective lens (as shown in Fig. 3.5), the near-focus amplitude h_A can be written as

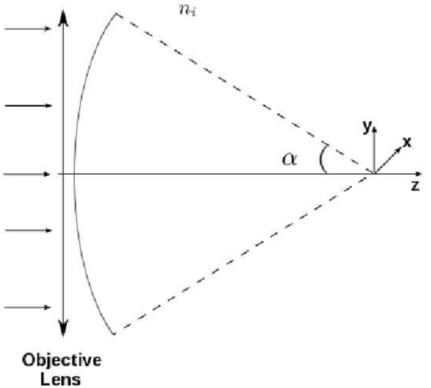


Figure 3.5: Schematic showing the focusing of an excitation light by an objective lens. α is the maximum semi-aperture angle and n_i is the RI of the objective immersion medium. The $\mathrm{NA} = n_i \sin\alpha$ is related to the amount of light entering the microscope (©Ariana-INRIA/I3S).

$$h_A(\mathbf{x}) = \int_{\mathbf{k}} \mathrm{OTF}_A(\mathbf{k}) \exp(j\mathbf{k}\cdot\mathbf{x})\, d\mathbf{k}\,. \tag{3.26}$$

By making the axial Fourier space co-ordinate k_z a function of lateral coordinates, $k_z = (k_i^2 - (k_x^2 + k_y^2))^{1/2}$, the computation of the PSF is reduced to N_z

number of 2D Fourier transforms

$$h_A(x,y,z) = \int_{k_x}\int_{k_y} P(k_x,k_y,z)\exp(j(k_x x + k_y y))dk_y dk_x\,, \tag{3.27}$$

where $P(\cdot,\cdot,\cdot)$ describes the overall complex field distribution in the pupil of a non-aberrated objective lens. The pupil function is a description of the magnitude and phase of the wavefront that a point source produces at the exit pupil of the imaging system. In simple terms, the above expression states that the field distribution at a point (x,y,z) in the image space can be obtained by applying Fourier transform on the overall pupil function. For a microscope suffering from defocus, the pupil function obtained by Hopkins [1955] can be written as

$$P(k_x,k_y,z) = \begin{cases} A(\theta_i)\exp(jk_0\varphi(\theta_i,\theta_s,z)), & \text{if } \frac{\left(k_x^2+k_y^2\right)^{\frac{1}{2}}}{k_i} < \frac{NA}{n_i}\,, \\ 0, & \text{otherwise,} \end{cases} \tag{3.28}$$

where $\text{NA} = n_i \sin\alpha$, $\varphi(\theta_i,\theta_s,z)$ is the optical phase difference between the wavefront emerging from the exit pupil and the reference sphere measured along the extreme ray, and $\theta_i = \arcsin(k_x^2+k_y^2)^{1/2}/k_i$ and $\theta_s = \arcsin(k_x^2+k_y^2)^{1/2}/k_s$ (*cf.* Fig. 3.6). θ_s and θ_i are related by Snell's law as

$$n_i \sin\theta_i = n_s \sin\theta_s\,. \tag{3.29}$$

The bandwidth of the pupil function here is $(2\pi\text{NA})/\lambda$.

The phase $\varphi(\theta_i,\theta_s,z)$ is a sum of two terms, the defocus term $\varphi_d(\theta_i,z)$ and the term due to aberrations $\varphi_a(\theta_i,\theta_s)$:

$$\varphi(\theta_i,\theta_s,z) = \varphi_d(\theta_i,z) + \varphi_a(\theta_i,\theta_s)\,. \tag{3.30}$$

For spherical aberration free imaging conditions, it is to be noted that $\theta_i \approx \theta_s$, and hence $\varphi_a(\theta_i,\theta_s) \approx 0$. This will be discussed in Subsection 3.4.1.

If f is the distance between the exit pupil and the geometrical image point, then Stokseth obtained the exact expression for this defect of focus as

$$\varphi_d(\theta_i,z;n_i) = -n_i\left(f + z\cos\theta_i - (f^2+2fz+z^2\cos^2\theta_i)^{\frac{1}{2}}\right)\,. \tag{3.31}$$

A small angle approximation can be obtained by making a series expansion of $\cos\theta_i$ and $\cos^2\theta_i$, and omitting terms of θ_i with higher order than two

$$\varphi_d(\theta_i,z;n_i)=\frac{1}{2}n_i z\theta_i^2\Big(1-\frac{z}{f+z}\Big),\theta_i\ll 1\,. \tag{3.32}$$

For small defocus and small angles, $z^2\cos^2\theta_i\approx z^2$, and hence Eq. (3.31) becomes

$$\varphi_d(\theta_i,z;n_i)\approx n_i z(1-\cos\theta_i)\,. \tag{3.33}$$

3.3.3 Equivalence in the Models

If we reconsider Eq. (3.26), we can rewrite $(\mathbf{k}\cdot\mathbf{x})$ in the spherical coordinates as

$$\mathbf{k}\cdot\mathbf{x}=k_i\sin\theta_i(x\cos\phi+y\sin\phi)\,, \tag{3.34}$$

and

$$d^2\mathbf{k}=k_i^2\sin\theta_i\cos\theta_i d\theta_i d\phi\,. \tag{3.35}$$

By substituting Eqs. (3.28), (3.33), (3.34), (3.35) into Eq. (3.27), we get

$$h_A(x,y,z)=k_i^2\int_0^\alpha\int_0^{2\pi}(\cos\theta_i)^{-\frac{1}{2}}\exp(jk_i\sin\theta_i(x\cos\phi+y\sin\phi))\times \\ \exp(jk_s z(1-\cos\theta_s))\sin\theta_i\cos\theta_i\, d\theta_i d\phi\,. \tag{3.36}$$

If $x=\rho\cos\psi$ and $y=\rho\sin\psi$, then $\rho=(x^2+y^2)^{1/2}$, and

$$h_A(x,y,z)=k_i^2\int_0^\alpha\int_0^{2\pi}(\cos\theta_i)^{\frac{1}{2}}\exp(jk_i\rho\sin\theta_i\cos(\phi-\psi))\times \\ \exp(jk_s z(1-\cos\theta_s))\sin\theta_i\, d\theta_i d\phi\,. \tag{3.37}$$

Remark 5. There are subtle differences between Eqs. (3.22) of Debye and (3.37) from Stokseth, namely

(a) the extra terms $\exp(jkz)$ and $\exp(j\pi/2)$ are canceled when the complex amplitude is squared to get the PSF, and

(b) the normalizing constant is different but the final intensities are re-normalized so that they both sum to unity.

3.3.4 Spatial Point-Spread Function Model

When we look at Eq. (2.33), theoretically speaking the estimation method on the object and PSF should be the same (Shepp & Vardi [1982]) as h and o play a symmetric role (Holmes [1992]; Biggs & Andrews [1997]). As was the case with object restoration, if there is no constraint imposed on the PSF, the solution is not unique. We saw at the beginning of Subsection 2.2 how a simple normalization on the PSF can remove the scaling ambiguity. Such a forced normalization ensures that the total energy of the PSF is bounded. Introducing constraints sequentially on the solution is a way of regularizing it and discarding other possibilities. In Holmes [1992], the constraints are introduces on the PSF so that the OTF lies within a specific volume. The PSF for the iterative algorithm is initialized by using the imaging parameters such as the NA, λ_{ex}, n_i, etc. It is difficult to know all these parameters as they are subject to changes during experimentation. The working NA of an objective lens, for example, cannot be known exactly. The PSF support is thus the maximum value it can take and does not reflect the true value. We will discuss more on this subject in the following chapter. A regularization model on the PSF (the cost function being $\mathcal{J}_{reg,h}(h)$) could also be argued along the same lines as the constraints introduced earlier for o. However, a TV kind of regularization on the PSF cannot model the continuity and regularity in the PSF. Such a ℓ_1 kind of norm on the PSF regularization is suitable only for PSFs that have edges, like motion blur (Chan & Wong [1998]). In such cases, the recovered PSF will be very much dependent on the object or specimen. Separation of the PSF and the object in this case becomes difficult as they have the same or similar solution space. We shall demonstrate this with an example in the next chapter. The regularization parameter, say λ_h, for such a model will be highly dependent on the amount of defocus, and will vary drastically from one image sample to another. To overcome these problems, the PSF could be intrinsically regularized through a parametric model. In this case the PSF estimation problem is reduced to estimating the parameters of the model.

Parametric Model

It is our contention that the best way to introduce constraints on the confocal PSF is in the parametric space. We thus assume that the unknown linear PSF belongs to a parametric family of functions $h(\boldsymbol{\omega}_h)$, where $\boldsymbol{\omega}_h \in \Theta \subseteq \mathbb{R}^d$, Θ denotes the space for all possible PSF parameters, and d is the number of free parameters. If the model is chosen well, the uncertainty lies in choosing the free parameters $\boldsymbol{\omega}_h$. By using the invariance property of ML estimation, we can say $\hat{h}_{ML}(\mathbf{x}) = h(\mathbf{x};\hat{\omega}_{h,\mathrm{ML}})$ is the MLE of the PSF. In a more general manner, any PSF can be written as the decomposition on a set of basis functions $\boldsymbol{\Phi}$ as

$$h(\mathbf{x}) \approx \sum_{l=1}^{N_h} w_l \Phi_l(\mathbf{x}) = \langle \mathbf{w}, \boldsymbol{\Phi}(\mathbf{x}) \rangle \,, \forall \mathbf{x} \in \Omega_s \,, \tag{3.38}$$

where w_l denotes the corresponding weights, $\langle \cdot, \cdot \rangle$ denotes vector dot product, and N_h is the number of basis components that approximate the PSF.

The distortions during image acquisition normally act as passive operations on the data, that is they neither absorb nor generate energy. Thus, when an object goes out of focus it is blurred, but the volume's total intensity remains constant. Consequently, all energy arising from a specific point in the fluorescent specimen should be preserved and

$$||h(\mathbf{x})||_1 = \sum_{\mathbf{x} \in \Omega_s} |h(\mathbf{x})| = 1 \,. \tag{3.39}$$

From Eq. (3.52), it is clear that the intensity distribution of a point source will always be positive and so

$$h(\mathbf{x}) \geq 0, \forall \mathbf{x} \in \Omega_s \,. \tag{3.40}$$

Finally, the PSF must satisfy an additional criterion of circular symmetry

$$h(-x,-y,z) = h(x,y,z), \forall (x,y,z) \in \Omega_s \,. \tag{3.41}$$

Radial symmetry is applicable only in cases where there is no additional aberrations due to the objective lens system and when it is horizontally well aligned. It is worthwhile to note that the constraints enforced via Eqs. (3.38)-(3.41) define convex subsets in the space of possible PSF solutions.

Separable Anisotropic Gaussian model One of the basis function that fits the above listed profile is the Gaussian kernel (*cf.* Santos & Young [2000] for the 2D case). If we assume aberration-free imaging, this model reduces the number of free parameters to estimate to just two. This simplistic model yet retains a reasonable fit to the actual PSF as was demonstrated by Zhang, et al. [2007] for a CLSM. It was found that a 3D separable Gaussian model gives a relative squared error (RSE)$<9\%$ for a pinhole diameter $D<3$AU and when the PSF peaks are matched (i.e. $||h(\mathbf{x})||_\infty = 1$), where we define the RSE as

$$\mathrm{RSE}(h_{\mathrm{Th}}, h) \overset{def}{=} \frac{||h_{\mathrm{Th}}(\mathbf{x}) - h(\mathbf{x})||_2^2}{||h_{\mathrm{Th}}(\mathbf{x})||_2^2}, \tag{3.42}$$

and $h_{\mathrm{Th}}(\mathbf{x})$ is the theoretical model given by Eq. (3.52). Thus, the diffraction-limited PSF (with restrictions on the pinhole diameter D) can be approximated as

$$h(\mathbf{x}) = (2\pi)^{-\frac{3}{2}} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right). \tag{3.43}$$

where $\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z)^T$ is the mean vector, $\Sigma = [\sigma_{ij}]_{1 \le i,j \le 3}$ is the covariance matrix, and $\mu_{(\cdot)}$, $\sigma_{(\cdot,\cdot)} \in \Theta$. As a first approximation, for thin specimens, the PSF is spatially zero centered, and $\boldsymbol{\mu} = \{0\}$. A mirror symmetry about the central xy-plane results in a diagonal covariance matrix and hence its determinant is $|\Sigma| = \sigma_r^4 \sigma_z^2$, where $\sigma_{11} (= \sigma_{22}) = \sigma_r$ and $\sigma_{33} = \sigma_z$ are the lateral and axial spreads in the image space. The unknown parameters that define the PSF are thus $\boldsymbol{\omega}_h = \{\sigma_r, \sigma_z\}$. It has to be mentioned that the PSF is anisotropic because, for microscopy, $\sigma_r \neq \sigma_z$ with $\sigma_z > \sigma_r$.

Nonseparable Gaussian model Another possible model that was proposed by Schlecht, et al. [2006], and that fits well with the theoretically derived expression in Eq. (3.52) is the following:

$$h(x,y,z) = (a_h)^{|z-\mu_z|} \left(2\pi\sigma_r(z)\right)^{-\frac{1}{2}} \exp\left(-\frac{\left(x-\mu_x\right)^2 + \left(y-\mu_y\right)^2}{2\sigma_r(z)} \right), (x,y,z) \in \Omega_s, \tag{3.44}$$

where $\sigma_r(z) = b_h|z - \mu_z| + c_h$, $\boldsymbol{\mu} = (\mu_x, \mu_y, \mu_z)^T$. We can safely assume that the radial means are centered i.e. $(\mu_x, \mu_y) = (0,0)$. The parameter of this function,

$\boldsymbol{\omega}_h$ are $\{a_h, b_h, c_h, \mu_z\}$. The base standard deviation is given by c_h, while b_h scales the distance from the central focal plane according to $(z - \mu_z) \in \mathbb{Z}$.

However, the disadvantage of using this above function is that it does not naturally sum to unity as in Eq. (3.39) and thus normalization has to be enforced. The other drawback is that this function is separable in the radial plane, but the lateral and axial planes are coupled. Finally, there are four parameters that define this nonseparable function, which is two more than the separable 3D case. In spite of these difficulties for Confocal PSF restoration, this model might be used in modeling and estimating the WFM PSF as it gives a 'better fit' (in the least squares sense) to the theoretical model.

Comments on the Approximations

- The parameters $\boldsymbol{\omega}_h$ in both Eqs. (3.43) and (3.44) are given in the image space. To obtain their values in the object space, they have to be normalized by the respective voxel sizes in the axial, Δ_{xy}, and radial space, Δ_z.
- If we consider the model in Eq. (3.43), it was shown by Zhang, et al. [2007] that the parameters $\boldsymbol{\omega}_h = \{\sigma_r, \sigma_z\}$ in the object space are dependent on the following settings: excitation wavelength λ_{ex}, RI n_i and the numerical aperture NA.
- From Eq. (3.52), we can intuitively infer that h_{Th} approaches a Gaussian distribution due to the 4^{th} power in the amplitude PSF h_{A} (if we momentarily assume the equality of excitation and emission light wavelengths) based only on the CLT. The Bessel function form of the h_A ensures that it is satisfied to a greater accuracy (*cf.* Fig. 3.8).
- In Fig. 3.8, we notice that the second power of the Bessel function has large side lobes and a wider principal maxima (in comparison to the fourth power). The WFM PSF, resembles this theoretical shape, and a separable Gaussian approximation would only lead to large model residues.

3.4 Aberrations in Fluorescence Microscopy

Under ideal conditions, a high NA objective lens transforms the planar wavefront incident on it to a spherical wavefront at the focal region. However, under practical situations, the refracted wavefront so produced has to go through several optical elements and through the specimen. Practical imaging systems are rarely aberration-free. Due to this reason, the emerging wavefront is rarely spherical in nature. Aberrant wavefront means that the resulting observed images will be distorted as well. While there are many aberrations that exist for the microscope, we restrict our analysis to spherical aberrations (SA), as this is the dominant and the most observable form.

3.4.1 Spherical Aberrations

SA is an optical effect occurring when the oblique rays entering a lens are focused in different locations than the central rays. SA is caused due to the mismatch between the RI of the objective lens immersion medium and the specimen embedding medium in fluorescence microscopy. When light crosses the boundary between materials with different refractive indices, it bends across the boundary surface differently depending on the angle of incidence (light refraction); the oblique rays are bent more than the central rays (*cf.* Török, et al. [1995]). If the mismatch is large, e.g. when going from oil lens immersion medium into a watery specimen embedding medium, SA causes the PSF to become asymmetric at depths of even a few microns. Also, the amount of light collected from a point source decreases with depth because of an axial broadening of the PSF. When measured experimentally using fluorophores (Shaevitz & Fletcher [2007]), it was found that the PSF changes from a fairly symmetric axial shape to an asymmetric shape. It is important to remember the following features of a depth varying PSF (DVPSF):

- in the absence of other aberrations, the PSF remains radially symmetrical;
- its peak intensity decreases with increase in the depth of focus, d[5];
- there is an increase in the FWHM of the PSF in the axial direction with increase in d;

---

[5] d is often referred to as the nominal focal position (NFP), and it is the depth of focus in a matched medium. It is also approximately equal to the paraxial focus.

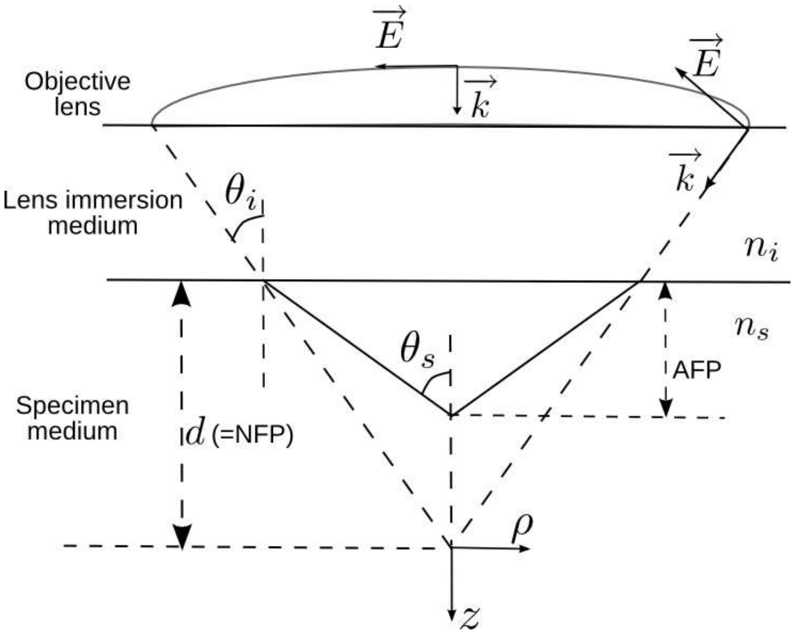


Figure 3.6: Schematic describing the focusing of rays when traveling between medium of different refractive indices. AFP is the actual focal position, and NFP is the nominal focal position in the absence of refractive index change. (ⓒAriana-INRIA/CNRS).

- if we ignore dispersion and absorption within the sample, the ℓ_1-norm of the total light intensity collected from a given object does not change with depth.

3.4.2 Pupil Phase Factor

We consider the situation where the objective lens focuses through an interface between media with different refractive indices as shown in Fig. 3.6. A point source is located at a depth d below the cover slip in a mounting medium of index n_s, observed with an objective lens designed for an immersion medium with a RI n_i (*cf.* Fig. 3.6). In Eq. (3.27), we had ignored the aberrations by assuming that $n_i \approx n_s$ and hence $\theta_i \approx \theta_s$. When $n_i \gg n_s$ (or $n_s \gg n_i$), we have to calculate the aberration function $\varphi_a(\theta_i, \theta_s)$, due to the mismatch of n_s and n_i. The phase change is determined by the difference between the optical path length traveled to the pupil by a ray that leaves the source at an angle θ_s relative to the optic axis and is refracted to the angle θ_i upon leaving the mounting medium, and the optical path length that a ray with angle θ_i would have traveled if the mounting medium index were the ideal index n_i. In their paper, Gibson & Lanni [1989] mention that there are 8 parameters (out of a total of 18) that may vary from their design conditions as recommended by the microscope manufacturer. However, when a microscope is properly calibrated, there are only 3 parameters that essentially vary under experimental conditions. These are

1. the specimen RI, n_s,
2. the immersion medium RI, n_i, of the objective lens, and
3. the depth, d, under the coverslip wherein lies the plane of focus.

The RI of the immersion medium is sensitive to changes in temperature especially if an oil immersion lens is used. An error in this parameter significantly affects the PSF calculated (*cf.* Hell & Stelzer [2006] and Juskaitis [2006]).

Using simple geometrical optics, we can show that this shift in the optical path can be analytically expressed as (*cf.* Born & Wolf [1999])

$$\varphi_a(\theta_i, \theta_s; d, n_i, n_s) = d(n_s \cos\theta_s - n_i \cos\theta_i). \tag{3.45}$$

The above phase term relies on the assumption that the error due to mismatch in the refractive indices between the cover glass n_g and the objective lens has either been compensated or is minimal. If the cover glass is used with an objective lens that is significantly different than its design specification, an additional phase term should be included, and d replaced by the thickness of the coverslip, t_g. In this case, the phase term would be

$$\varphi_a(\theta_g, \theta_s; d, n_i, n_s) = t_g(n_g \cos\theta_g - n_i \cos\theta_i) + d(n_s \cos\theta_s - n_g \cos\theta_g), \tag{3.46}$$

where n_g is the RI of the cover glass, and θ_g is the angle relative to the optic axis to which light is refracted in the cover glass.

Actual microscope objectives are only corrected to remove SA at specific wavelengths usually, but not necessarily to the same wavelengths for which they are chromatically corrected. Immersion oils used have not only a specific RI but also a specific dispersion (*cf.* Hell & Stelzer [2006] for variation of the RI with wavelength; a phenomenon known as dispersion). The lens from a specific manufacturer require both to be corrected.

3.4.3 Point-Spread Function Approximations

In Subsection 3.3.4, we introduced a spatial approximation for the diffraction-limited PSF, and in this section, we propose an approximate for the PSF using the pupil phase.

Nonlinear Phase Approximation

The phase term in Eq. (3.45) can be rewritten as

$$\begin{aligned}\varphi_a(\theta_i,\theta_s;d,n_i,n_s) &= d\sec\theta_i(n_s\cos\theta_s\cos\theta_i - n_i\cos^2\theta_i)\,,\\ &= d\sec\theta_i(n_s\cos\theta_s\cos\theta_i - n_i(1-\sin^2\theta_i))\,,\end{aligned} \tag{3.47}$$

as $\cos^2\theta_i = (1-\sin^2\theta_i)$ and $\sec\theta_i = 1/\cos\theta_i$. From Snell's law, $n_s\sin\theta_s = n_i\sin\theta_i$ and from basic trigonometry we know that $\cos(\theta_s-\theta_i) = (\cos\theta_s\cos\theta_i + \sin\theta_s\sin\theta_i)$. Thus,

$$\begin{aligned}\varphi_a(\theta_i,\theta_s;d,n_i,n_s) &= d\sec\theta_i(n_s\cos\theta_s\cos\theta_i - n_i + n_s\sin\theta_s\sin\theta_i)\,,\\ &= d\sec\theta_i\left(n_s\left(\cos\theta_s\cos\theta_i\sin\theta_s\sin\theta_i\right) - n_i\right)\,,\\ &= d\sec\theta_i(n_s\cos(\theta_s-\theta_i) - n_i)\,.\end{aligned} \tag{3.48}$$

If we assume that the difference is not very large, $\theta_s \approx \theta_i$, so that

$$\varphi_a(\theta_i,\theta_s;d,n_i,n_s) \approx d\sec\theta_i(n_s - n_i)\,. \tag{3.49}$$

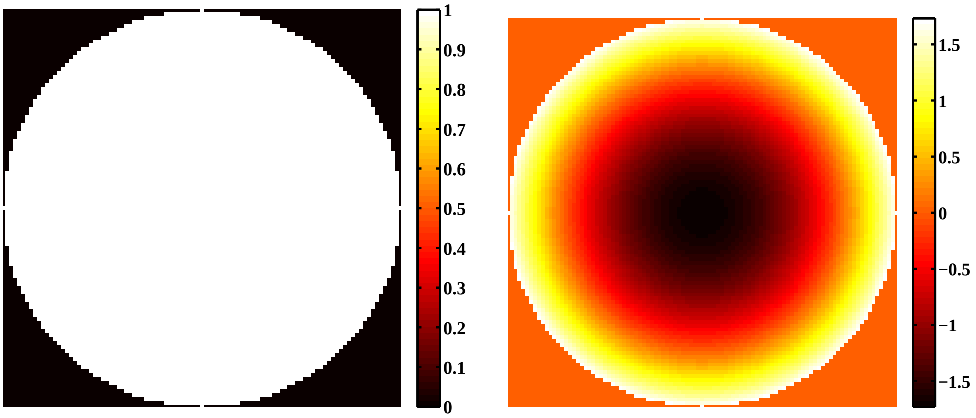
With this nonlinear phase approximation, we can say that the phase is dependent only on the depth d and the difference in the indices $(n_s - n_i)$.

Linear Phase Approximation

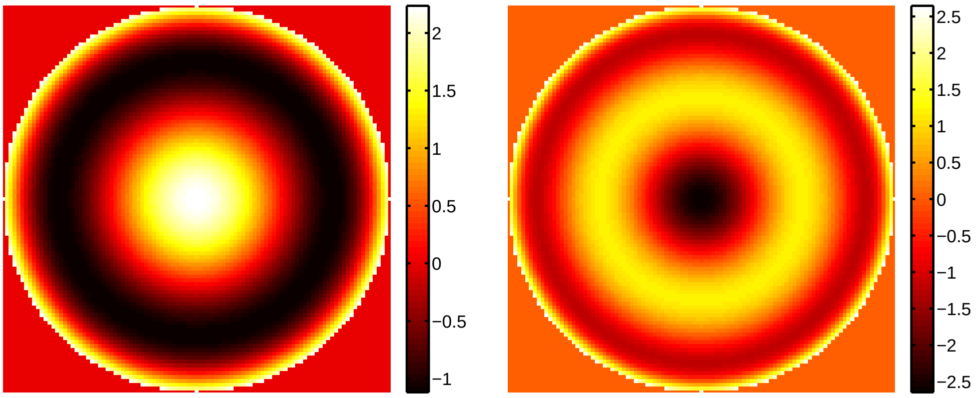
The phase term in Eq. (3.30) can be approximated by a using a set of circular basis functions called Zernike polynomials (Booth, et al. [1998]). These polynomials form a complete orthogonal set on the unit disk. If the microscope is calibrated and if the objective lens correctly chosen, there is no azimuthal variation, and it is sufficient to consider Zernike circle polynomials of order n and zero kind $(Z_n^0(\mathbf{k}_r;d,n_i))$. Here $\mathbf{k}_r = (k_x^2 + k_y^2)^{1/2}$ is the normalized radial coordinate in the pupil plane such that $0 \leq \mathbf{k}_r \leq 1$. The phase could be expanded using these polynomials as

$$\varphi_a(\mathbf{k}_r; d, n_i, n_s) = d\,\mathrm{NA}\left(\sum_{n=0}^{4} c_n Z_n^0(\mathbf{k}_r; d, n_i, n_s)\right), \tag{3.50}$$

where the expansion coefficients (c_n) needs to be calculated or determined. In Dieterlen, et al. [2004], these polynomials were used to reconstruct an experimentally obtained but noisy PSF. However, as our interest lies in reconstructing the object, we are interested only in defocus and the first order SA. That is, $\varphi(\mathbf{k}_r; d, n_i, n_s)$ is approximated to only the Zernikes of up to 4th order. In Fig. 3.7,



(a) Constant or Piston $n=0$, (b) defocus $n=2$,



(c) primary SA $n=4$ (d) secondary SA $n=6$

Figure 3.7: Zernike polynomials for the pupil phases with zero azimuthal frequency and radial orders (a) Zero (Piston), (b) two (defocus), (c) four (primary spherical aberration), (d) six (secondary spherical aberration) (ⓒAriana-INRIA/I3S).

we show the different Zernike polynomials for orders zero, two, four and six respectively. We recall that for the PSF expression derived in Section 3.3.2, as there are no aberrations, only the Zernikes up to 2nd order need to be calculated. Although the secondary SA for order $n = 6$ is not common in fluorescence microscopy, it is shown here as comparison with the primary SA.

Apodization Function Approximation

For the gradient calculations in Subsection 5.2.2, we make an approximation on the apodization function $A(\theta_i)$ of Eq. (3.28). We simply assume the apodization function for excitation and emission are the same and

$$A(\theta_i) \approx 1\,. \tag{3.51}$$

Using the above expression, the apodization resembles the ideal pupil in Fig. 3.4(a) or the Piston in Fig. 3.7(a).

3.5 Three Dimensional Incoherent Point-Spread Function

In this section we summarize the theory that has been discussed earlier for obtaining the theoretical PSF expression. In numerical calculations of the PSF, we will primarily use the Stokseth approximation. For the diffraction-limited case, Eq. (3.27) gives the complex amplitude PSF, with the pupil function as given in Eq. (3.28) and the defocus phase in Eq. (3.33). While for the SA case, an additional phase term should also be included as given by Eq. (3.45). These expressions are simple to implement using Fourier transforms. For a 1p excitation, when the fluorophore returns to the ground state, by the *Stoke's shift*, the emitted wavelength is longer than the excitation wavelength. From the Helmholtz reciprocity theorem, the analytical PSF expression for a CLSM is written as a product of the excitation distribution and the emission distribution as

$$h_{\mathrm{Th}}(\mathbf{x};\lambda_{\mathrm{ex}},\lambda_{\mathrm{em}}) = C\,|h_A(\mathbf{x};\lambda_{\mathrm{ex}})|^2 \times \iint_{\Sigma_p} |h_A(x-x_1, y-y_1, z;\lambda_{\mathrm{em}})|^2 dx_1 dy_1\,, \tag{3.52}$$

where λ_{ex} and λ_{em} are the excitation and emission wavelengths respectively, $\Sigma_p = \{(x_1, y_1) : x_1^2 + y_1^2 \leq D^2/4\}$ characterizes the circular pinhole, D is the back-projected diameter of the circular pinhole, and C is a scaling constant. From Eq. (3.52), it is easy to see that the PSF satisfies the following properties:

$$h_{\text{Th}}(\mathbf{x}) \geq 0, \forall \mathbf{x} \in \Omega_s, \tag{3.53}$$

and bounded as

$$\int_{\mathbf{x} \in \Omega_s} h_{\text{Th}}(\mathbf{x}) d\mathbf{x} < \infty . \tag{3.54}$$

In literature, the PSF expressions are primarily defined using the Bessel functions, and we will derive them beginning with the following definitions:

Definition 6. Bessel function of first kind and integer order n

$$J_n(x) = \frac{1}{\pi} \int_0^{\pi} \cos(n\theta - x \sin\theta)\, d\theta, \forall x \in \mathbb{R}, n \in \mathbb{N}, \tag{3.55}$$

Corollary 3.5.1. *Bessel function of first kind and zero order*

$$J_0(x) = \frac{1}{\pi} \int_0^{\pi} \cos(x \sin\theta)\, d\theta \forall x \in \mathbb{R}, \tag{3.56}$$

In Fig. 3.8, we consider $J_0(x)$, and its higher powers (two and four). We notice that as the function's power is raised, the side lobes become smaller. The full width at half maximums (FWHM) of the central maxima also decreases. This property was exploited in Subsection 3.3.4 to model spatial PSF approximation for CLSM.

Eq. (3.27) can be modified to now include the additional phase term as

$$h_A(\mathbf{x}; \lambda) = C \int_0^{\alpha} A(\theta_i) \sin\theta_i J_0(k_i \rho \sin\theta_i) \exp(j k_0 \varphi_d(\theta_s, z; n_i)) \times \\ \exp(j k_0 \varphi_a(\theta_i, \theta_s; d, n_i, n_s)) d\theta_i . \tag{3.57}$$

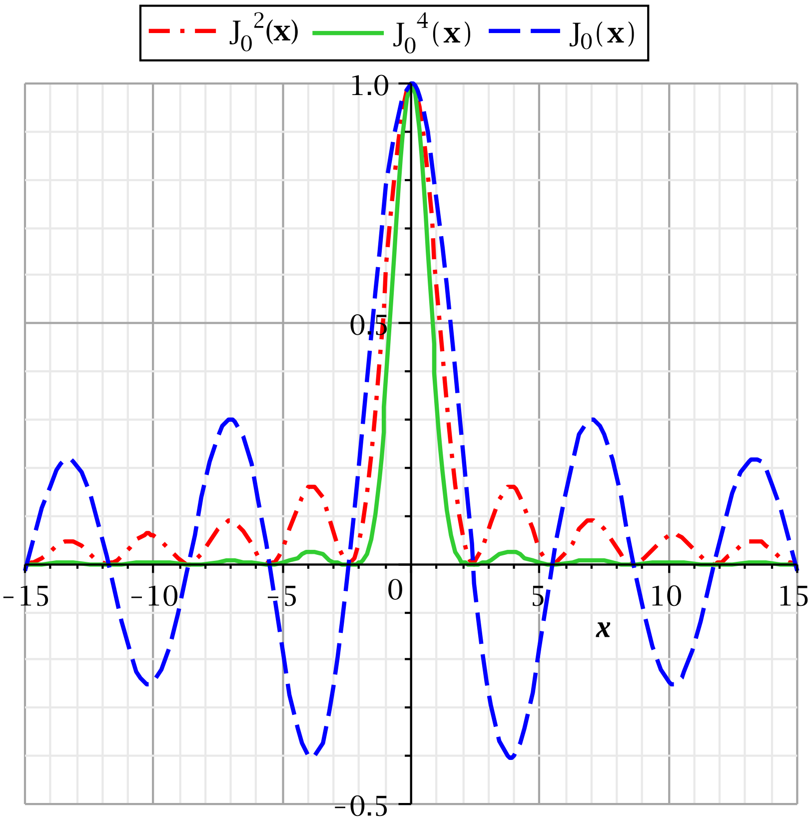


Figure 3.8: Plot of the Bessel function of first kind with order zero. As the function's power is raised, the side lobes becomes smaller and so does the FWHM of the central maxima.

and

$$h_A(\mathbf{x};\lambda)=2\pi k_i^2 \exp(jk_s z)\, C \int_0^{\alpha} (\cos\theta_i)^{\frac{1}{2}} \sin\theta_i J_0(k_i \rho \sin\theta_i) \times \exp(jk_0 \varphi_a(\theta_i,\theta_s;d,n_i,n_s)) \exp(-jk_s z \cos\theta_s)\, d\theta_i \,. \tag{3.58}$$

By using a normalized radial co-ordinate $t=\sin\theta_i/\sin\alpha$, the field at $\mathscr{P}_0$ becomes

$$h_A(\mathbf{x};\lambda)=2\pi k_i^2 \exp(jk_s z)\, C \int_0^{1} (1-t^2\sin^2\alpha)^{-\frac{1}{2}}\, t \sin^2\alpha J_0(k t \rho \sin\alpha) \times \exp(jk_0 \varphi_a(t;d,n_i,n_s)) \exp(-jk_s z(1-t^2\sin^2\alpha)^{\frac{1}{2}})\, dt \,. \tag{3.59}$$

For small angles (*i.e.* NA≤ 0.7), $\sin\theta_i \approx \theta_i$ and hence

$$h_A(\mathbf{x};\lambda) = 2\pi k_i^2 \sin\alpha\, C \int_0^1 J_0(k t \rho \sin\alpha) \exp(j k_0 \varphi_a(t; d, n_i, n_s)) t\, dt\,, \tag{3.60}$$

and the quadratic aberration phase term is

$$\varphi_a(t; d, n_i, n_s) = d\left((n_s^2 - t^2 n_i^2 \sin^2\alpha)^{\frac{1}{2}} - (n_i^2 - t^2 n_i^2 \sin^2\alpha)^{\frac{1}{2}}\right). \tag{3.61}$$

Definition 7. For a circularly symmetrical function $f(\rho,\theta) \equiv f(\rho)$, the *Fourier-Bessel* transform or *Hankel* transform of zero order is

$$\mathscr{B}[f(\rho)] = 2\pi \int_0^\infty \rho f(\rho) J_0(2\pi\rho k)\, d\rho\,. \tag{3.62}$$

Thus, the Fourier transform of a circularly symmetric function is itself circularly symmetrical.

Definition 8. We define a circle function as

$$\mathrm{circ}(\rho) = \begin{cases} 1, & \rho < 1 \\ \frac{1}{2}, & \rho = 1 \\ 0, & \text{otherwise.} \end{cases} \tag{3.63}$$

Corollary 3.5.2. *The Fourier-Bessel transform of the circle function can be written as*

$$\mathscr{B}[circ(\rho)] = 2\pi \int_0^1 \rho J_0(2\pi\rho k)\, d\rho\,. \tag{3.64}$$

Using a change of variable $\rho' = 2\pi\rho k$ *and the identity* $\int_0^x \xi J_0(\xi) d\xi = x J_1(x)$,

$$\mathscr{B}[circ(\rho)] = \frac{1}{2\pi k^2} \int_0^{2\pi k} \rho' J_0(\rho') d r' = \frac{1}{k} J_1(2\pi k)\,, \tag{3.65}$$

where $J_1(\cdot)$ *is a Bessel function of the first kind and order one.*

By using the notation of Fourier-Bessel transform or Hankel transform of zero-order, Eq. (3.60) is simplified as

$$h_A(\mathbf{x};\lambda) = C k_i^2 \sin^2\alpha\, \mathscr{B}\left[\exp(j k_0 \varphi_a(t;d,n_i,n_s))\right] . \qquad (3.66)$$

Numerically Computing the Point-Spread Function

From Eqs. (3.33) and (3.45), the phase of the refracted wavefront could be computed, and the 2D pupil function derived for a fixed defocusing. The 2D PSF is calculated as the fast fourier transform (FFT) of the 2D pupil. The 3D PSF is then reconstructed as stacks of the 2D PSF. This numerical computation is applicable to calculate the PSF for most fluorescent microscopes. For a CLSM, the resulting emission PSF is convolved with a pinhole model. This method requires in total $2 \times N_z$ number of 2D FFTs, and two 3D FFTs for the pinhole convolution.

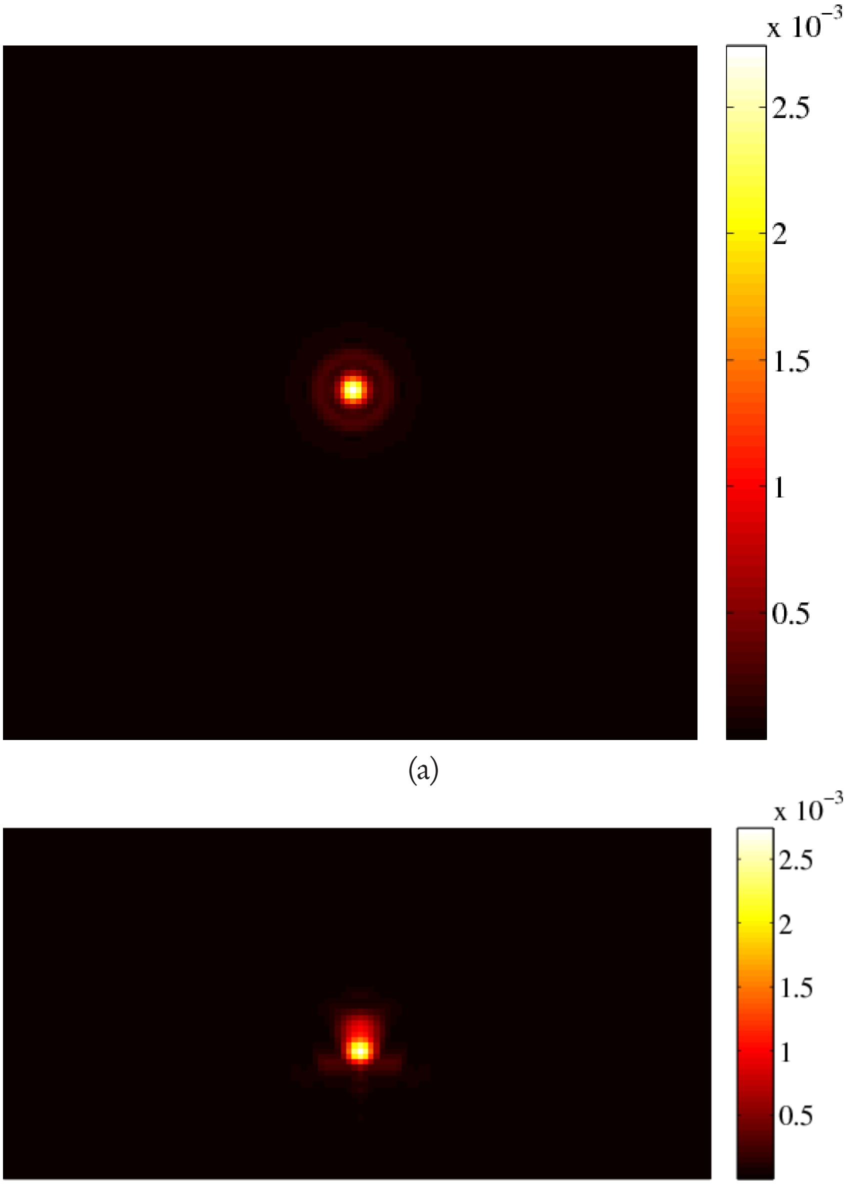
Fig. 3.9 shows a numerically computed confocal PSF by using the above technique and Eq. (3.52). The microscope uses an excitation beam with a peak wavelength $\lambda_{\text{ex}} = 488$nm, and the emission peak has a wavelength $\lambda_{\text{em}} = 520$nm. The objective lens is is an immersion oil type ($n_i = 1.518$) Plan-Neofluar, with a NA of 1.3, and 40X magnification. The depth d under the coverslip is 15μm and into a specimen immersed in water ($n_s = 1.33$).

3.5.1 Validity of the Scalar Model

In this theoretical derivation of PSF expressions, the most important approximation used is the treatment of light as a scalar phenomenon. This neglects the fundamentally vectorial nature of the electromagnetic fields. The scalar theory yields accurate results if two conditions are met:

- the diffraction aperture Σ is very large in comparison to the wavelength of light λ, and
- the diffracting fields must not be observed too close to the aperture *i.e.* $r_{01} \gg \lambda$, $r_{21} \gg \lambda$.

It is important to mention that for all microscopes operating in the far-field region, the above approximations are justified.



(a)

(b)

Figure 3.9: Maximum intensity projection (MIP) of a numerically computed spherically aberrated confocal PSF (a) along the optic axis giving the radial plane, and (b) along lateral plane giving the axial plane (ⓒAriana-INRIA/I3S).

3.6 Conclusion

In this chapter, we derived an analytical expression for the PSF based on Stokseth's model. The model was extended to include spherical aberrations by modifying the phase of the pupil function. The aberrated pupil phase was derived from geometrical optics. The proposed nonlinear aberrated phase approximation shows that it essentially depends on the depth d and the difference $(n_s - n_i)$. Linear approximations for the pupil phase based on Zernike polynomials were also provided. These models could be numerically implemented by applying FFTs on the complex pupil function. This analytical PSF model will be used in the next chapters to generate a simulated observation and also for image restoration.

CHAPTER 4

Blind Deconvolution for Thin Specimens

"It was six men of Hindustan to learning much inclined,
Who went to see the Elephant (Though all of them were blind)
That each by observation might satisfy the mind ...
And so these men of Hindustan disputed loud and long,
Each in his own opinion exceeding stiff and strong,
Though each was partly in the right and all were in the wrong ..."[6]

-John Godfrey Saxe (American Poet)

We propose an alternate minimization (AM) iterative algorithm for estimating the parameters of the confocal laser scanning microscope's (CLSM) point-spread function (PSF), and the specimen fluorescence distribution. If the problem of deconvolution is ill-posed, that of blind deconvolution is under-determined as the number of unknowns to be estimated is increased without any increase in the input observation data. To make the deconvolution problem well-posed, at every iteration, the estimated object intensity is regularized using a total variational potential function. A new prior is suggested that ensures positivity of the intensity estimation. As direct restoration from the observation data is very difficult, an underlying model for the PSF is defined as well. An anisotropic, separable 3D Gaussian model is used to restrict the PSF solution space. The object estimation method is treated in Subsection 4.2.1 and the PSF estimation in Subsection 4.2.2. In Subsection 4.2.3, we propose a method to handle the parameter of the regularization model. We provide numerical simulations, in Section 4.3,

---

[6] *The Blindmen and the Elephant*, a poem based on a Buddhist fable, found in the *Udana*, chapter 6, section 4.

to support the Bayesian framework, and also demonstrate on fluorescent microspheres and real data. The results from the phantom images show that the PSF can be estimated to a high degree of accuracy, and those on real data show better deconvolution as compared to a full theoretical PSF model. We conclude in Section 4.4 with a discussion and proposed future work. The scope of this chapter is restricted to restoring images from a CLSM given the spatial invariance nature of the diffraction-limited PSF.

4.1 Introduction

As the PSF describes the response of the microscope to every point in a sample, restoration by deconvolution could be achieved only if this information is available. The *non-blind* case involves using either the PSF obtained experimentally (McNally, et al. [1994]; Shaevitz & Fletcher [2007]) by imaging a small fluorescent bead (so as to approximate a point object) positioned just under the cover slip, or by numerically calculating it. Although an empirically obtained PSF should be an ideal choice for a deconvolution algorithm, it suffers from low contrast, its FWHM is much larger than theoretically expected, and the images are contaminated by noise and aberrations (*cf.* Chapter 5). On the contrary, analytical model of the PSF (Stokseth [1969]; Gibson & Lanni [1989]) are free from noise and takes into account the acquisition system's physical information as parameters. This information however might not be available or might change during the course of the experiment (for example, due to heating of live samples). It was observed by Preza, et al. [1992] that an inaccurate PSF can affect the performance of the deconvolution algorithm. We hence arrive at the blind approach of estimating both the specimen and the unknown PSF. When imaging biological cells, the amount of radiation must be limited. Plant cells are especially attuned to light collection and do not have any protection against flux levels of $100 - 10{,}000$ times higher than the normal levels (*cf.* Shaw [2006]). The amount of incident light should be controlled to avoid cellular damage and photo bleaching. Thus, there is often only a single observation of the specimen volume available for restoration. The non-repeatability of the experiments makes it difficult to restore the lost frequencies beyond the diffraction limits unless some information about the object or the PSF is provided. In the absence of any information on the object or the PSF, restoration from the observation data will be like the deductions of the blind men in Saxe's poem. In this chapter, we propose a BD algorithm that estimates the diffraction-limited PSF (under aberration free conditions) and use it to restore the true object fluorescence. The aberrations can be ignored iff the NFP$< 15\mu$m, and so the method proposed in this chapter is applicable only to thin samples where the effect of diffraction is dominant over aberrations. The extension to thick specimens is discussed in Chapter 6.

4.2 Blind Deconvolution by Alternate Minimization

As mentioned before, the simultaneous estimation of o and h directly from the observed images is a very difficult task. At the beginning of Subsection 2.2, an AM algorithm was summarized in Eqs. (2.26) and (2.27). This algorithm alternatively estimates the o first while keeping the PSF function h constant, and then update the PSF using the previous object estimate. In this section, we propose the AM approach by making use of the proposed prior object model in Eq. (2.21) and PSF model in Eq. (3.43).

4.2.1 Estimation of the Object

For the time being let us assume that either the PSF or its parameters $\omega_h \in \Theta$ are known (either by initialization or from previous estimates) and hence $\hat{h}$ is determinate. By substituting Eqs. (2.21) and (2.33) in Eq. (2.24), we get

$$\Pr(o,\hat{h}|i) = Z_{\lambda_o}^{-1} \exp\left(-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|\right) \times \prod_{\mathbf{x}\in\Omega_s} \frac{\left(\left(\hat{h}*o+b\right)(\mathbf{x})\right)^{i(\mathbf{x})} \exp\left(-\left(\hat{h}*o+b\right)(\mathbf{x})\right)}{i(\mathbf{x})!}. \quad (4.1)$$

As in Eq. (2.26), by applying $-\log$ operator to the *a posteriori* above, the cost function $\mathcal{J}(o,\hat{h}|i)$ to be minimized with respect to o becomes

$$\mathcal{J}(o,\hat{h}|i)=\sum_{\mathbf{x}\in\Omega_s}\left(\hat{h}*o+b\right)(\mathbf{x})-\sum_{\mathbf{x}\in\Omega_s}i(\mathbf{x})\log\left(\hat{h}*o+b\right)(\mathbf{x})+ \lambda_o\sum_{\mathbf{x}\in\Omega_s}|\nabla o(\mathbf{x})|+\log\left(Z_{\lambda_o}\right). \quad (4.2)$$

If we choose h to be a Dirac, then the minimization of the above energy function gives a denoised image while a non-Dirac PSF leads to simultaneous deconvolution and denoising. There are many approaches proposed in literature to minimize the above cost function. In the following subsection, we will discuss one such method based on the EM algorithm with TV regularization.

Expectation Maximization Algorithm with Total Variation Regularization

The Euler-Lagrange equation for minimizing $\mathcal{J}(o,\hat{h}|i)$, given by Eq. (4.2), with respect to o is

$$1-\hat{h}(-\mathbf{x})*\left(\frac{i(\mathbf{x})}{\left(\hat{h}*o+b\right)(\mathbf{x})}\right)-\lambda_o\operatorname{div}\left(\frac{\nabla o(\mathbf{x})}{|\nabla o(\mathbf{x})|}\right)=0, \quad (4.3)$$

where $\hat{h}(-\mathbf{x})$ is the Hermitian adjoint operation on $\hat{h}(\mathbf{x})$ and div stands for the divergence operator (*cf.* Dey, et al. [2006] for details). Eq. (4.3) can be solved for the object o by adopting the MLEM algorithm (*cf.* Appendix B) with TV regularization. We adopt an explicit scheme that proposes to minimize Eq. (4.2) by the following multiplicative algorithm:

$$\hat{o}^{(n+1)}(\mathbf{x})=\left(\frac{i(\mathbf{x})}{\left(\hat{o}^{(n)}*\hat{h}+b\right)(\mathbf{x})}*\hat{h}(-\mathbf{x})\right)\cdot\frac{\hat{o}^{(n)}(\mathbf{x})}{1-\lambda_o\operatorname{div}\left(\frac{\nabla\hat{o}^{(n)}(\mathbf{x})}{|\nabla\hat{o}^{(n)}(\mathbf{x})|}\right)} \quad (4.4)$$

where $(\cdot)$ denotes the Hadamard multiplication (component wise) and n the iteration number for the deconvolution algorithm. We stop the deconvolution algorithm if the difference measure between two successive iterations is smaller than a specified threshold ϵ. The term $\operatorname{div}\left(\nabla\hat{o}^{(n)}(\mathbf{x})/|\nabla\hat{o}^{(n)}(\mathbf{x})|\right)$ can be numerically implemented with the use of central differences and the *minmod* scheme (*cf.* Appendix D).

Positivity and Flux Constraints

The deconvolution algorithm that was described above suffers from an inherent weakness. For large values of λ_o, even when the starting guess $\hat{o}^{(n)}$ (with $n = 0$) is positive, the successive estimates need not necessarily have positive intensities. We know that the true intensity of the object $o(\mathbf{x})$ is always non-negative. Most algorithms often truncate these negative intensities to zero or to a small positive value. This however is a crude manner to handle the estimated intensities as it can lead to loss of essential information and sometimes also introduce bias into the calculations.

So how else can the problems associated with negative intensity estimates be handled? Fortunately, the problem is entirely due to the poor statistical methodology adopted. The modification that we suggest is to include this knowledge of non-negative true intensities into the prior term of Eq. (2.21). The distribution that would express precisely this condition is

$$\Pr[o(\mathbf{x})] \sim \begin{cases} Z_{\lambda_o}^{-1} \exp\left(-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|\right), & \text{if } o(\mathbf{x}) \geq 0 \\ 0, & \text{otherwise.} \end{cases} \tag{4.5}$$

For the sake of numerical differentiability, we approximate the prior Eq. (4.5) using a *sigmoid* function as

$$\Pr[o(\mathbf{x})] \sim Z_{\text{new},\lambda_o}^{-1} \exp\left(-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|\right) \left(\frac{1}{1+\exp\left(\beta_o(\epsilon - o(\mathbf{x}))\right)}\right), \tag{4.6}$$

where ϵ is a small value close to zero, and β_o is a value that specifies the steepness of the *sigmoid* function. Typically the values of β_o and ϵ are chosen to be very large and small respectively as precision might allow. Their values do not individually affect the algorithm and hence need not be known accurately. Fig. 4.1 shows the variation in the sigmoid curve as a function of the steepness factor β_o. ϵ is kept fixed constant throughout and is close to zero. We see that as the value of

β_o increases, the Sigmoid curve approaches a hard-limiting function (at $o(\mathbf{x})=0$) while at $\beta_o=0$ it is an all-pass function.

The cost function in Eq. (4.2) becomes,

$$\mathcal{J}(o,\hat{h}|i)=\sum_{\mathbf{x}\in\Omega_s}\left(\hat{h}*o+b\right)(\mathbf{x})-\sum_{\mathbf{x}\in\Omega_s}i(\mathbf{x})\log\left(\hat{h}*o+b\right)(\mathbf{x})+$$

$$\lambda_o\sum_{\mathbf{x}\in\Omega_s}|\nabla o(\mathbf{x})|+\log\left(Z_{\text{new},\lambda_o}\right)+\log\left(1+\exp(\beta_o(\epsilon-o(\mathbf{x})))\right), \quad (4.7)$$

the Euler-Lagrange equation in Eq. (4.3) is now

$$1-\hat{h}(-\mathbf{x})\left(\frac{i(\mathbf{x})}{(h*o+b)(\mathbf{x})}\right)-\lambda_o\,\text{div}\left(\frac{\nabla o(\mathbf{x})}{|\nabla o(\mathbf{x})|}\right)-$$

$$\beta_o\frac{\exp(\beta_o(\epsilon-o(\mathbf{x})))}{1+\exp(\beta_o(\epsilon-o(\mathbf{x})))}=0, \quad (4.8)$$

and the multiplicative algorithm of Eq. (4.4) becomes

$$\hat{o}^{(n+1)}(\mathbf{x})=\left[\frac{i(\mathbf{x})}{\left(\left[\hat{o}^{(n)}*\hat{h}\right]+b\right)(\mathbf{x})}*\hat{h}(-\mathbf{x})\right]\cdot$$

$$\frac{\hat{o}^{(n)}(\mathbf{x})}{1-\lambda_o\,\text{div}\left(\frac{\nabla\hat{o}^{(n)}(\mathbf{x})}{|\nabla\hat{o}^{(n)}(\mathbf{x})|}\right)-\beta_o\frac{\exp(\beta_o(\epsilon-o(\mathbf{x})))}{1+\exp(\beta_o(\epsilon-o(\mathbf{x})))}}. \quad (4.9)$$

Intuitively, we see that the cost function given by Eq. (4.7) ensures that the energy for negative intensity pixels ($o(\mathbf{x})<\epsilon$) are very high and hence are not reachable (or not possible solutions) during the iteration procedure.

If the PSF is normalized and in the absence of a background signal, it is easy to show that for each iteration of the MLEM algorithm (Eq. (4.4) with $\lambda_o=0$) the following property is true: $e=\sum_{\mathbf{x}\in\Omega_s}i(\mathbf{x})=\sum_{\mathbf{x}\in\Omega_s}\hat{o}(\mathbf{x})$. This property is known as the flux or global photometry conservation and it guarantees that the total number of counts of the reconstructed object is the same as the total number of observation counts. However, this property is lost with regularization and can be incorporated by modifying the cost function Eq. (4.2) to an additive form similar to Bratsolis & Sigelle [2001] or by enforcing it in the following manner

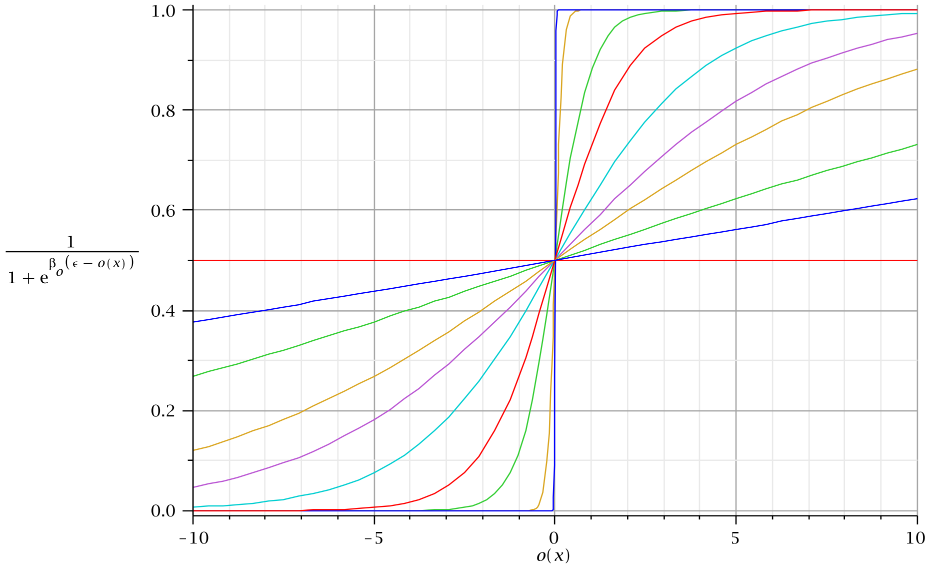


Figure 4.1: The function $1/(1+\exp(\beta_o(\epsilon-o(\mathbf{x}))))$ in the new prior $\Pr(o(\mathbf{x}))$ as a function of the object function $o(\mathbf{x})$ for different steepness factors β_o (©Ariana-INRIA/I3S).

after every iteration:

$$\hat{o}^{(n+1)}_{\text{new}}(\mathbf{x})=\left(\frac{e^{(0)}}{e^{(n+1)}}\right)\hat{o}^{(n+1)}_{\text{old}}(\mathbf{x}), \tag{4.10}$$

where $e^{(n+1)}=\sum_{\mathbf{x}\in\Omega_s}\hat{o}^{(n+1)}_{\text{old}}(\mathbf{x})$, and $e^{(0)}=\sum_{\mathbf{x}\in\Omega_s}\hat{o}^{(0)}(\mathbf{x})=\sum_{\mathbf{x}\in\Omega_s}i(\mathbf{x})$.

4.2.2 Point-Spread Function Parameter Estimation

The method outlined in Subsection 4.2.1 requires the knowledge of the PSF $\hat{h}(\mathbf{x})$ or $h(\mathbf{x};\hat{\omega}_h)$. If we re-examine Eq. (2.27) in the light of the parametric model proposed in Section 3.3.4, we can say $\hat{h}_{\text{MAP}}(\mathbf{x})\approx h(\mathbf{x},\hat{\omega}_{h,\text{MAP}})$, and

$$\hat{\omega}_{h,\text{MAP}}=\underset{\omega_h>0}{\operatorname{argmax}}\{\Pr(i|\hat{o},h(\omega_h))\Pr(\omega_h)\}\,, \tag{4.11}$$

where $\Pr(\omega_h)$ is the prior on the parameters ω_h. If we assume that the parameters are uniformly distributed in a set $\left[\omega_{h,\text{LB}},\omega_{h,\text{UB}}\right]$, then $\Pr(\omega_h)=1/(\omega_{h,\text{UB}}-\omega_{h,\text{LB}})$, $\forall\omega_h\in\left[\omega_{h,\text{LB}},\omega_{h,\text{UB}}\right]$. We define the set $\left[\omega_{h,\text{LB}},\omega_{h,\text{UB}}\right]$ by $\mathscr{S}$. From Eqs. (2.33), (4.11) and with the invariance property of ML estimate, minimizing the energy function with respect to the PSF ($\mathcal{J}(\hat{o},h|i)$) or the parameters ($\mathcal{J}(\hat{o},\omega_h|i)$) are equivalent. Thus,

$$\mathcal{J}(\hat{o},\omega_h|i) = -\sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x})\log\left(h(\omega_h)*\hat{o}\right)(\mathbf{x}) + \sum_{\mathbf{x}\in\Omega_s}\left(h(\omega_h)*\hat{o}\right)(\mathbf{x}). \tag{4.12}$$

If the true object o is assumed to be known *a priori* as $\hat{o}$, then estimation of the parameters of the PSF is straight forward as the cost function in Eq. (4.12) is convex in the neighborhood of the optimal $\omega_h \in \mathcal{S}$. This can be demonstrated by using a simple simulation. In Fig. 4.2, the cost function $\mathcal{J}(\omega_h, o|i)$ is plotted when the object o and the observation i are known. As we are in the simulation case, the true parameters are known as well. However, what we show by this toy problem is that, even when one of the parameters is known with an error of ±11%, the other parameter could be estimated accurately. Since, the cost function is locally convex about the optimal parameter value, we can hence use a gradient-descent (GD) kind of algorithm (*cf.* Atkinson [1989]) for estimating the parameters. Analytically minimizing Eq. (4.12) with respect to the parameters leads us to the following:

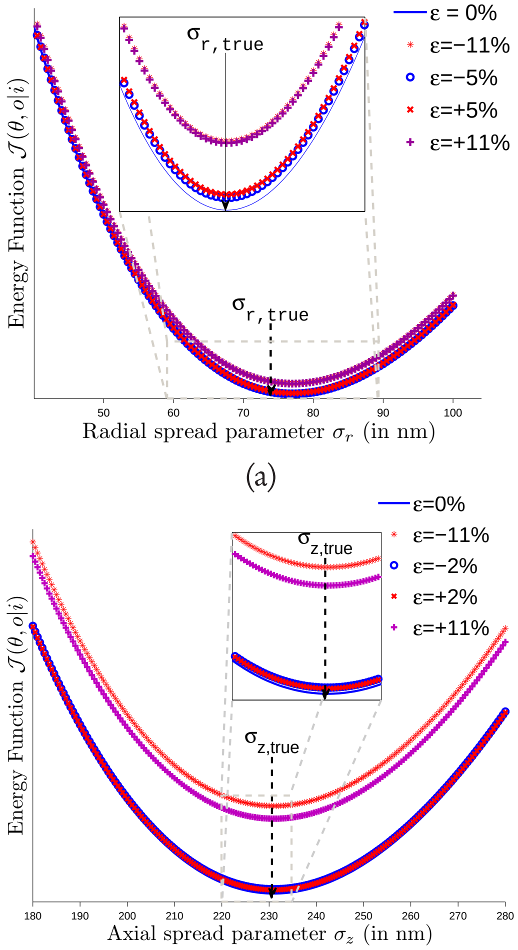
$$\hat{\omega}_h^{(n+1)} = \hat{\omega}_h^{(n)} - \tau^{(n)}\nabla_{\omega_h}\mathcal{J}\left(\hat{o},\hat{\omega}_h^{(n)}|i\right), \forall\hat{\omega}_h \in \mathcal{S}, \tag{4.13}$$

where $\tau^{(n)}$ and $\nabla_{\omega_h}\mathcal{J}(\hat{o},\hat{\omega}_h^{(n)}|i)$ are the step size and the search directions at iteration n. The gradient of the cost function in Eq. (4.12) with respect to the parameters can be calculated by using the chain rule as

$$\begin{aligned}\nabla_{\theta_l}\mathcal{J}(\hat{o},\omega_h|i) = \sum_{\mathbf{x}\in\Omega_s} \frac{\partial}{\partial\omega_h}\left(h(\omega_h)*\hat{o}\right)(\mathbf{x}) - \\ \sum_{\mathbf{x}\in\Omega_s} \frac{i(\mathbf{x})}{\left(h(\omega_h)*\hat{o}\right)(\mathbf{x})}\frac{\partial}{\partial\omega_h}\left(h(\omega_h)*\hat{o}\right)(\mathbf{x}),\end{aligned} \tag{4.14}$$

where

$$\left[\frac{\partial}{\partial\omega_h}h(\omega_h)\right]_{\omega_h=\sigma_r} = \left(-\frac{2}{\sigma_r} + \frac{x^2+y^2}{\sigma_r^3}\right)h(\omega_h), \tag{4.15}$$



(a)

(b)

Figure 4.2: Variation of the energy function $\mathcal{J}(o, \omega_h|i)$ with respect to (a) lateral parameter σ_r, and (b) axial PSF parameter σ_z. For this simulation, the true object o is known and the observation is generated using a known 3D Gaussian model. The axial PSF parameter σ_z is varied by a factor $\pm\epsilon$ to monitor its effect on the estimated lateral parameter σ_r and vice versa. The $\sigma_{(r,\text{true})}$ and $\sigma_{(z,\text{true})}$ shown are the true parameter values (©Ariana-INRIA/I3S).

and

$$\left[\frac{\partial}{\partial \omega_h} h(\omega_h)\right]_{\omega_h=\sigma_z} = \left(-\frac{1}{\sigma_z} + \frac{z^2}{\sigma_z^3}\right) h(\omega_h). \tag{4.16}$$

The separable nature of the Gaussian distribution reduces the complexity of the algorithm, as the convolution with the 3D Gaussian PSF can be implemented as three successive 1D multiplications in the separable Fourier domain. Only a single DFT of the object estimate $\hat{o}$ needs to be performed as an analytical closed form expression for the Fourier transform of the Gaussian and its derivative exists and can be numerically calculated (*cf.* Appendix C). The GD computation is terminated if the difference measure between two successive iterations is smaller than ϵ (in practice 10^{-3} or 10^{-4}). The estimated parameters from Eq. (4.13), can now be used to numerically calculate the PSF by substituting the estimates into the Eq. (3.43) after converting them into the image space parameters.

Theoretical Parameter Calculations

For faster convergence of the estimation algorithm, we calculate the theoretical limits for the 3D separable Gaussian model based on the work of Zhang, et al. [2007]. If the PSF models are supposed to be diffraction-limited and aberrations ignored, the 3D lateral Gaussian spread for the non-paraxial case (*i.e.* $\mathrm{NA} > 0.7$) can be computed as

$$\hat{\sigma}_{r,\mathrm{np}}^{\mathrm{clsm}} = \frac{0.421\lambda_{\mathrm{ex}}\lambda_{\mathrm{em}}}{n_i\left(\lambda_{\mathrm{ex}}^2+\lambda_{\mathrm{em}}^2\right)^{1/2}}\left[\frac{\cos^{3/2}\alpha-1}{7\cos^{3/2}\alpha-3\cos^{7/2}\alpha-4}\right]^{\frac{1}{2}}. \tag{4.17}$$

Similarly, the 3D axial Gaussian parameter for the non-paraxial case is

$$\sigma_{z,\mathrm{np}}^{\mathrm{clsm}} = \frac{0.86\lambda_{\mathrm{ex}}\lambda_{\mathrm{em}}}{n_i\left(\lambda_{\mathrm{ex}}^2+\lambda_{\mathrm{em}}^2\right)^{1/2}}\times \left[\frac{\left(\cos^{3/2}\alpha-1\right)^2}{4\cos^5\alpha-25\cos^{7/2}\alpha+42\cos^{5/2}\alpha-25\cos^{3/2}\alpha+4}\right]^{\frac{1}{2}}. \tag{4.18}$$

For the paraxial case ($\mathrm{NA} \leq 0.7$), the lateral Gaussian parameter for 3D PSF is

$$\hat{\sigma}_{r,p}^{\mathrm{clsm}} = \frac{0.225\lambda_{\mathrm{ex}}\lambda_{\mathrm{em}}}{\mathrm{NA}\left(\lambda_{\mathrm{ex}}^2+\lambda_{\mathrm{em}}^2\right)^{1/2}}, \tag{4.19}$$

and the axial Gaussian parameter

$$\hat{\sigma}_{z,p}^{\mathrm{clsm}} = \frac{0.78 n_i\lambda_{\mathrm{ex}}\lambda_{\mathrm{em}}}{\mathrm{NA}^2\left(\lambda_{\mathrm{ex}}^2+\lambda_{\mathrm{em}}^2\right)^{1/2}}. \tag{4.20}$$

Using l'hôpital's rule, it was shown by Zhang, et al. [2007] that when $\mathrm{NA} = n_i \sin\alpha$ approaches zero, the nonparaxial parameters Eqs. (4.17) and (4.18) tend to the paraxial parameters Eqs. (4.19) and (4.20).

Theoretical Parameter Limits

By using Eqs. (4.17) and (4.18), we numerically determined the range of values that the PSF parameters could possibly take by evaluating them for 96 different objective lens. These objectives were of immersion types dry, water, glycerol and oil, with NA ranging from 0.025 (1× dry Plan-Neofluar lens) to 1.46 (100× Plan-Apochromat oil immersion lens). We found that for an excitation source of wavelength $\lambda_{\mathrm{ex}} = 488$nm and emission of $\lambda_{\mathrm{em}} = 520$nm, $\sigma_r \in [71, 4394]$nm and $\sigma_z \in [128, 446703]$ nm.

4.2.3 Regularization Parameter Handling

A delicate situation is in the choice of the regularization parameter λ_o; too small values yield overly oscillatory estimates owing to noise or discontinuities, while too large values yield overly smooth estimates. The selection or estimation of the regularization parameter is thus a critical issue on which there are several proposed approaches (for example *cf.* Jalobeanu, et al. [2002]). However, we are looking for a simple technique that could be combined with the AM algorithm and also fits well with the Bayesian framework. The difficulty in performing marginalization with respect to λ_o is that the partition function is not easily computed. An approach to circumvent this problem is to approximate the partition function $Z_{\mathrm{new},\lambda_o}$ as suggested by Mohammad-Djafari [1996]. Accordingly, we say that

$$\Pr(o|\lambda_o) \sim \frac{1}{Z_{\mathrm{new},\lambda_o}} \exp\left(-\lambda_o \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|\right)\left(\frac{1}{1+\exp\left(\beta_o(\epsilon - o(\mathbf{x}))\right)}\right), \quad (4.21)$$

with the partition function approximated as

$$Z_{\mathrm{new},\lambda_o} \approx \lambda_o^{\alpha_{\lambda_o} N_x N_y N_z}, \quad (4.22)$$

where α_{λ_o} is a parameter of the partition function. The full posterior probability density function in Eq. 2.24 now becomes

$$\Pr(o,h,\lambda_o|i)=\frac{\Pr(i|o,h)\Pr(h)\Pr(o|\lambda_o)\Pr(\lambda_o)}{\Pr(i)}, \tag{4.23}$$

and we define the hyperprior $\Pr(\lambda_o)$ using the Gamma distribution $\lambda_o \sim \Gamma(a_g,b_g)$ as

$$\Pr(\lambda_o)\sim \lambda_o^{a_g-1}\exp\left(-b_g\lambda_o\right), \tag{4.24}$$

where a_g is the shape and b_g is the scale hyperparameters for the hyperprior. The expectation for this distribution is given by $\mathbb{E}(\lambda_o)=a_g/b_g$, and variance $\mathbb{V}(\lambda_o)=a_g/b_g^2$. In Fig. 4.3, we see that when the hyperparameters (a_g,b_g) are $(1,0)$, the hyperprior in Eq. (4.24) becomes the uninformative uniform distribution. Most of the work on this subject is based on this model, and it was also used in Pankajakshan, et al. [2009b] as well. On the other hand, when $(a_g,b_g)=(0,0)$, Eq. (4.24) leads to a noninformative Jeffreys hyperprior. Using the hyperprior

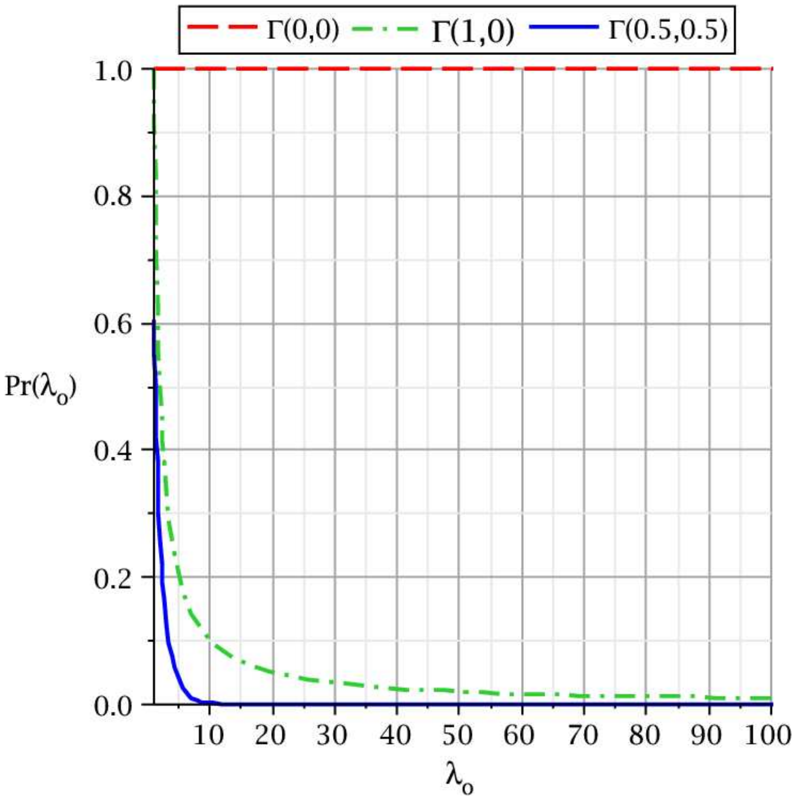


Figure 4.3: Gamma distribution for three different pairs of shape and scaling parameters (ⓒAriana-INRIA/I3S).

expression in Eq. (4.24), the cost function can be written in terms of the cologarithm of Eq. (4.23) as

$$\mathcal{J}(o,h,\lambda_o|i)=\mathcal{J}_{\text{obs}}(i|o,h)+\mathcal{J}_{\text{reg},h}(h)+\mathcal{J}_{\text{reg},o}(o|\lambda_o)+\mathcal{J}_{\text{hp}}(\lambda_o), \tag{4.25}$$

where $\mathcal{J}_{\text{reg},o}(o|\lambda_o)=-\log\Pr(o|\lambda_o)$ and $\mathcal{J}_{\text{hp}}(\lambda_o)=-\log\Pr(\lambda_o)$. The regularization parameter could be estimated by

$$\hat{\lambda}_o=\underset{\lambda_o\geq 0}{\operatorname{argmin}}\,\mathcal{J}_{\text{reg}}(o,h,\lambda_o|i) \tag{4.26}$$

$$=\underset{\lambda_o\geq 0}{\operatorname{argmin}}\,\mathcal{J}_{\text{reg}}(o|\lambda_o)+\mathcal{J}_{\text{hp}}(\lambda_o), \tag{4.27}$$

By substituting Eqs. (4.21) and (4.24) in Eq. (4.25),

$$\begin{aligned}\mathcal{J}(o,h,\lambda_o|i)=&\mathcal{J}_{\text{obs}}(i|o,h)+\mathcal{J}_{\text{reg},h}(h)-\\ &\log\left(\lambda_o^{-\alpha_{\lambda_o}N_xN_yN_z}\exp\left(-\lambda_o\sum_{\mathbf{x}}|\nabla o(\mathbf{x})|\right)\right)-\\ &\log\left(\lambda_o^{a_g-1}\exp\left(-b_g\lambda_o\right)\right),\end{aligned} \tag{4.28}$$

and

$$\begin{aligned}\mathcal{J}(o,h,\lambda_o|i)=&\mathcal{J}_{\text{obs}}(i|o,h)+\mathcal{J}_{\text{reg},h}(h)+(\alpha_{\lambda_o}N_xN_yN_z-a_g+1)\log\lambda_o+\\ &\lambda_o\sum_{\mathbf{x}}|\nabla o(\mathbf{x})|+b_g\lambda_o.\end{aligned} \tag{4.29}$$

We get the gradient of $\mathcal{J}(o,h,\lambda_o|i)$ with respect to λ_o, from Eqs. (4.27) and (4.29), by getting rid of the terms independent of λ_o as

$$\begin{aligned}\frac{\partial}{\partial\lambda_o}\mathcal{J}(o,h,\lambda_o|i)=&\frac{\partial}{\partial\lambda_o}\Big(\left(\alpha_{\lambda_o}N_xN_yN_z-(a_g-1)\right)\log\lambda_o+\lambda_o\sum_{\mathbf{x}}|\nabla o(\mathbf{x})|+\\ &b_g\lambda_o\Big),\end{aligned} \tag{4.30}$$

$$\frac{\partial}{\partial \lambda_o}\mathscr{J}(o,h,\lambda_o|i)=\left(\frac{\alpha_{\lambda_o}N_xN_yN_z-a_g+1}{\lambda_o}+\sum_{\mathbf{x}}|\nabla o(\mathbf{x})|+b_g\right). \tag{4.31}$$

The optimum value, $\hat{\lambda}_o$, is a solution of the following

$$\left(\frac{\alpha_{\lambda_o}N_xN_yN_z-a_g+1}{\lambda_o}+\sum_{\mathbf{x}}|\nabla o(\mathbf{x})|+b_g\right)=0,\forall\lambda_o\geq 0. \tag{4.32}$$

Since the object $o(\mathbf{x})$ is unknown, we substitute it with its estimate $\hat{o}(\mathbf{x})$,

$$\hat{\lambda}_o=\left(\frac{a_g-1-\alpha_{\lambda_o}N_xN_yN_z}{\sum_{\mathbf{x}}|\nabla\hat{o}(\mathbf{x})|+b_g}\right),\ \forall a_g>\left(1+\alpha_{\lambda_o}N_xN_yN_z\right). \tag{4.33}$$

4.3 Results

In this section, we validate the proposed AM algorithm on both synthetic and real data.

4.3.1 Algorithm Analysis

The global procedure alternatively minimizes the cost function in Eq. (4.2) first with respect to o in Eq. (4.4) while keeping the PSF function h fixed, and then updates the PSF Eq. (4.13) using the previous object estimate $\hat{o}$. Since the iterative algorithm requires an initial guess for the true object, we use the mean of the observed image (i.e. every site is assumed to have a constant intensity and is equally likely) for the initialization. As there are no constraints on the PSF or its support, initialization of the parameters to small values cannot guarantee its convergence to the desired size (due to the Dirac trivial solution). To avoid this problem, we choose the initial parameters to be utmost $2\varkappa^{-1}$ Resels[7] and $9\varkappa^{-1}$ Resels for the lateral and axial case respectively, and descend down to the optimal value.

Both $\mathscr{J}_{\text{obs}}(i|o,h)$ and $\mathscr{J}_{\text{reg},o}(o)$ in Eq. (4.7) are convex though not in the strict

---

[7] 1 Resel$=0.61\lambda_{\text{ex}}/\text{NA}$; $\varkappa=2.35$

Algorithm 2: The Proposed Blind Deconvolution Algorithm.

```
begin
    Input: Observed volume $i(\mathbf{x})$, $\forall \in \Omega_s$.
    Data: Initial parameters $\hat{\omega}_h^{(0)} \in \mathscr{S}$ (cf. Subsection 4.3.1),
          convergence criterion $\epsilon$, parameters $(a_g, b_g, \alpha_{\lambda_o})$.
    Output: Deconvolved volume $\hat{o}(\mathbf{x})$, PSF parameters
            $\hat{\omega}_h \in \mathscr{S} \subset \mathbb{R}_+^2$.
    Initialization: $n \leftarrow 0$, $\hat{o}^{(n)}(\mathbf{x}) \leftarrow \texttt{Mean}(i(\mathbf{x}))$, $\hat{h}^{(n)}(\mathbf{x}) \leftarrow h(\mathbf{x}; \hat{\omega}_h^{(n)})$
    Eq. (3.43).
    $\lambda_o$ estimation: $\hat{\lambda}_o \leftarrow (a_g - 1 - \alpha_{\lambda_o})/(\sum_{\mathbf{x}}(|\nabla \hat{o}^{(0)}(\mathbf{x})|) + b_g)$ from
    Eq. (4.33).
    Estimate the background $\hat{b}$ once (Subsection 1.3.2).
    while $|\hat{\omega}_h^{(n)} - \hat{\omega}_h^{(n-1)}|/\hat{\omega}_h^{(n)} \geq \epsilon$ do
        Using the minmod scheme in Appendix D, calculate
        $\mathrm{div}(\nabla \hat{o}^n(\mathbf{x})/|\nabla \hat{o}^n(\mathbf{x})|)$.
        Deconvolution: Calculate $\hat{o}^{(n+1)}$ from Eq. (4.9).
        Projection Operation: Scale $\hat{o}^{(n+1)}$ for flux preservation
        Eq. (4.10).
        Parameter estimation: Calculate $\hat{\omega}_h^{(n+1)}$ from Eqs. (4.13)
        and (4.14).
        Assign: $\hat{h}^{(n+1)}(\mathbf{x}) \leftarrow h(\mathbf{x}; \hat{\omega}_h^{(n+1)})$ and $n \leftarrow (n+1)$.
    end
end
```

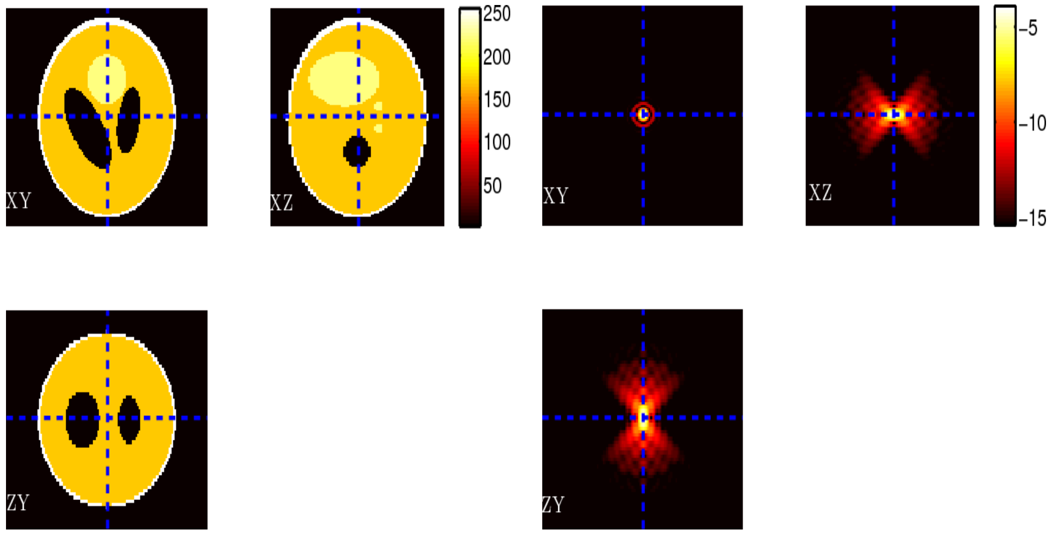
sense. Although the convergence of the algorithm to the optimal solution is theoretically difficult to prove, numerical experiments indicate that the global procedure does converge when the initialization is carried out as described above. When the gradient step α_n in Eq. (4.13) is well chosen, the cost function $\mathscr{J}$ decreases at every iteration on o and ω_h so that $\mathscr{J}(o^{n+1}, \omega_h^{n+1}|i) < \mathscr{J}(o^n, \omega_h^n|i)$ and $\mathscr{J}$ is inferior bounded.

4.3.2 Experiments on Simulated Data

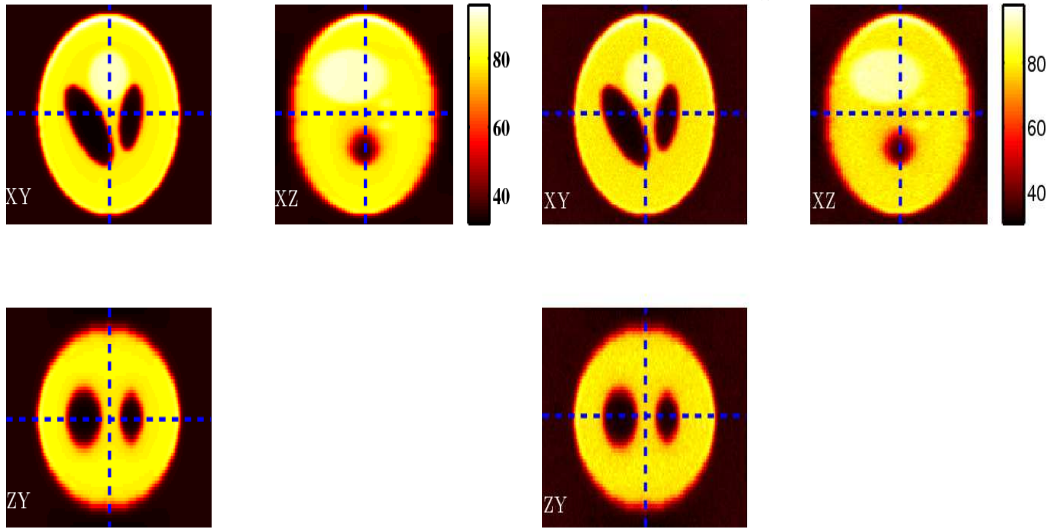
In this subsection, we present some results on simulated data. Performing experiments on such a data set permits us to both quantitatively and qualitatively measure the improvement over the observation, and also compare the deconvolution result with the original object. Since the PSF in this case is available, we can also verify the effectiveness of the BD algorithms by estimating it.

Fig. 4.4(a) shows a 3D phantom test object of dimension $128 \times 128 \times 64$, with XY and Z pixels sampled at 20nm and 50nm respectively. The blurred images in Fig. 4.4(c) were generated by blurring the object using an analytically modeled PSF as in Fig. 4.4(b). A uniform level of 10 IU was added as an average intensity to account for background and stray fluorescence. The noise is assumed to be shot noise arising due to low-photon imagery, modeled as Poisson statistics with $\gamma = 100$. For generating the PSF, the excitation λ_{ex} and emission λ_{em} peak wavelengths were chosen as 488nm and 520nm. The objective is a *C-Apochromat* lens with a NA of 1.2 and 63× magnification. The objective immersion medium is water, $n_i = 1.33$, and imaging is assumed to be done at a depth $d = 0\mu$m. The diameter of the pinhole was assumed to be vanishingly small.

In Fig. 4.5(a) we show the normalized gradient $\nabla_x(i(\mathbf{x}))/|\nabla_x(i(\mathbf{x}))|$ of the observation in Fig. 4.4(d), but along the x direction. While the divergence of the normalized gradient is shown in Fig. 4.5(b). The vector field flow lines for the gradient of the observation is shown along two directions in Fig. 4.6(a) and (b). The contour levels are that of the original object, and it is displayed for visualizing the difference in flow lines in homogeneous regions and edges. The size of the arrows give the amplitude of the field at the point in the volume, and the direction. We notice that in the homogeneous regions, the size of the arrows are very small. When no noise is present, they vanish completely. On the other hand, at the edges, the field magnitude is large and so are the arrow lengths. In Fig. 4.7, we compare the results of the restoration with other blind approaches that was explained in Subsection 2.2. For visual comparison, we re-normalized all the intensities to the levels of the simulated object. The stopping threshold ϵ and the regularization parameter λ_o, dictates the amount of reduction in the blur and noise-edge effects respectively. The stopping threshold ϵ, between two successive iterations was fixed as 10^{-4}. The regularization parameter λ_o is estimated once from $\hat{o}^{(0)}(\mathbf{x})$, and is fixed constant for the rest of the iterations. This ensures that the calculations are at a manageable level. We chose the parameters $(a_g, b_g, \alpha_{\lambda_o})$ to be $(0.5, 0.5, -1)$ (*cf.* Fig. 4.3). The result of the proposed approach is shown in Fig. 4.7(b) with the hyperparameter $\hat{\lambda}_o = 0.0097$. For comparisons, we chose to



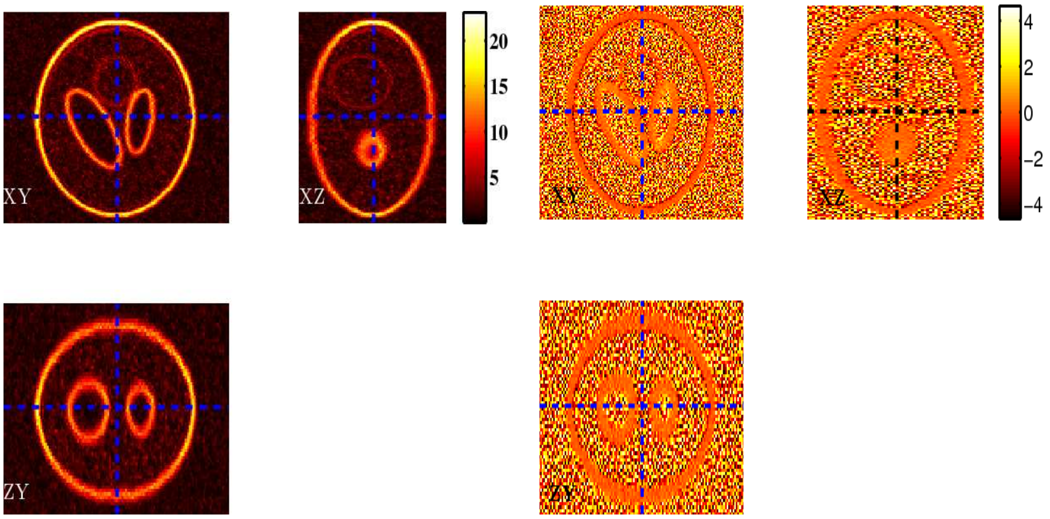
(a) Object volume, (b) point-spread function,



(c) blurred object, (d) observed volume

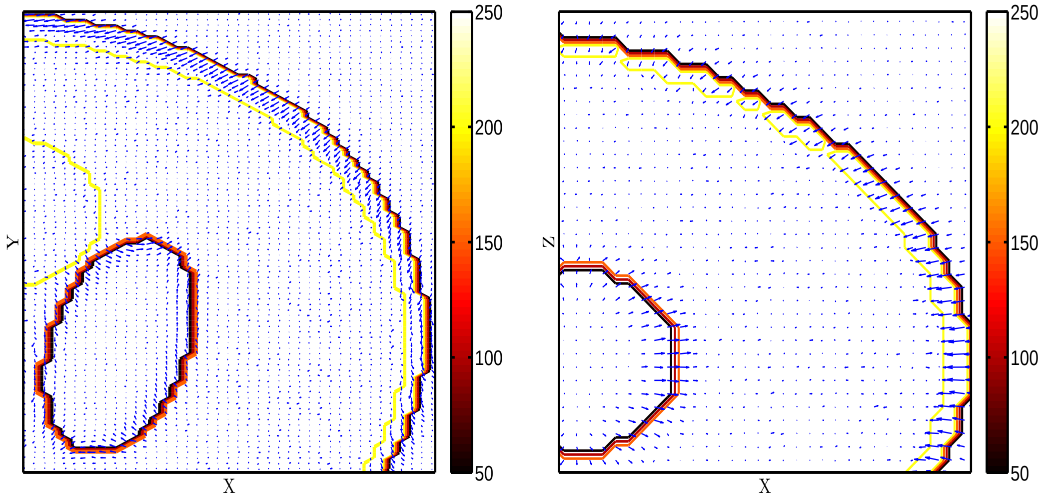
Figure 4.4: Numerical simulation of an object, point-spread function, blurred object and observation. The pixel size is 20nm in the radial direction and with a step of 50nm in the axial directions. The number of photon counts per pixel, γ, is assumed to be 100. While the PSF is shown using a logarithmic contrast stretch, the object and the observation are shown on a linear scale (ⓒAriana-INRIA/I3S).

accelerate naïve MLEM BD algorithm, with the q_o parameter in Eq. (2.34) as 1.4. The iteration was terminated at 30 in Fig. 4.7(c) by making a visual trade-off between deconvolution and ringing artifacts. The PSF solutions were band-limited as in Holmes [1992]. If the algorithm was continued beyond 30 iterations, we began to notice the noise artifacts discussed also in the reference of Holmes [1992].



(a) (b)

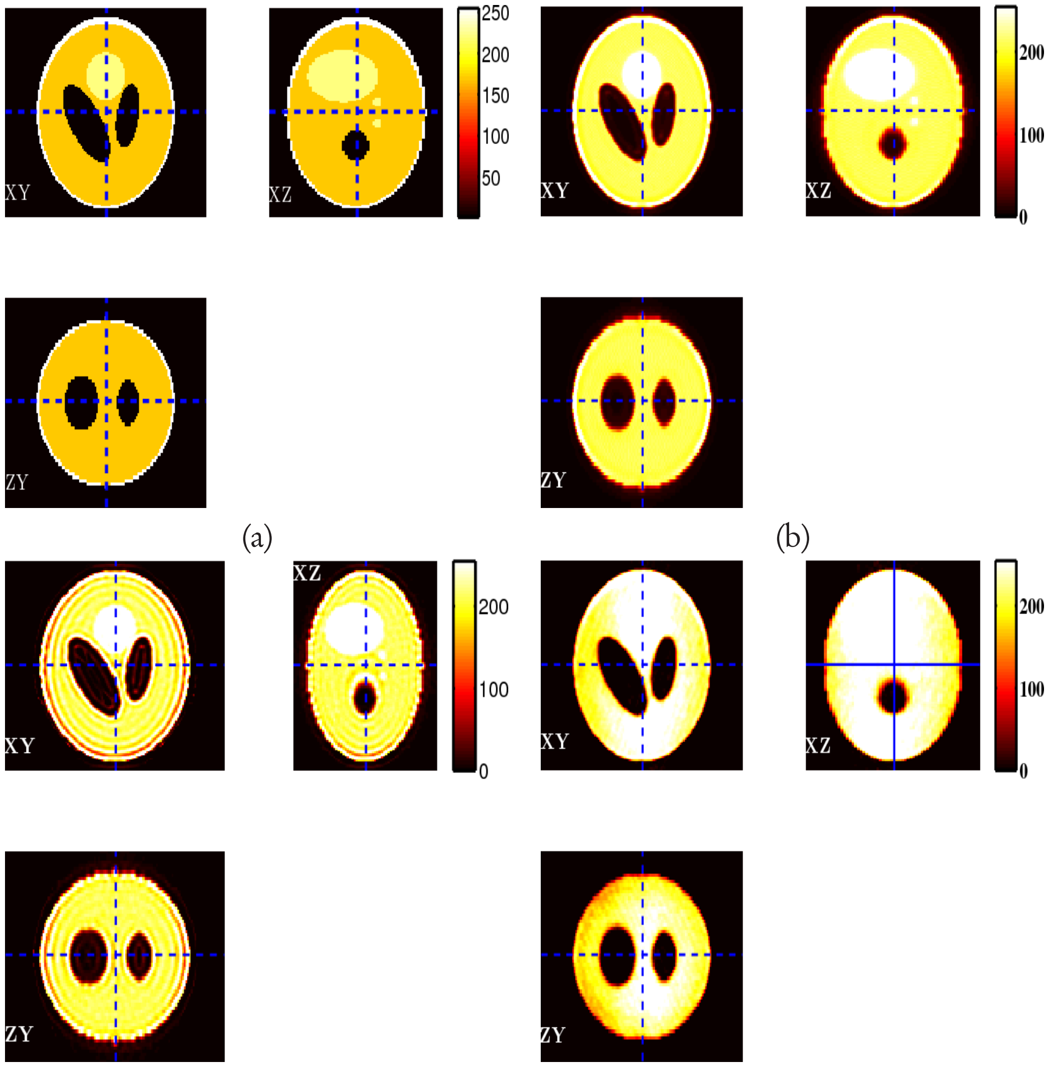
Figure 4.5: (a) The normalized gradient of a simulated observation $\nabla_x(i(\mathbf{x}))/|\nabla_x(i(\mathbf{x}))|$ along x direction, and its (b) total variation function $\mathrm{div}\left(\nabla_x(i(\mathbf{x}))/|\nabla_x(i(\mathbf{x}))|\right)$ (ⓒAriana-INRIA/I3S).



(a) (b)

Figure 4.6: The gradient vector field flow lines along (a) xy plane and (b) yz plane. The contours are shown to note the difference between the homogeneous regions from the edges (ⓒ Ariana-INRIA/I3S).

We observed that for the IBD approach (Ayers & Dainty [1988]), a good initialization is very important. Hence, we chose a 3D median filtered observation as the initial estimate and hand picked the parameters of the algorithm for best results. The algorithm was accelerated by averaging the successive estimates with an acceleration parameter of 0.7, the noise thresholding parameter as 0.1, and the



(a) (b)

(c) (d)

Figure 4.7: (a) Original 3D object, and the blind restoration results using (b) proposed approach, (c) an accelerated naïve MLEM approach at 30 iterations (Holmes [1992]) and (d) IBD (Ayers & Dainty [1988]) with the parameters given in the text (©Ariana-INRIA/I3S).

maximum number of iterations as 4. The initial PSF was obtained by calculating the parameters as in Subsection 4.2.2. As the deconvolution results quickly diverged beyond the 4th iteration, the algorithm was manually terminated, and the result at the 4th iteration is displayed in Fig. 4.7(d). It was found that our proposed approach estimated the PSF much closer (with minimum RSE) to the true analytically modeled PSF used to simulate the observations. The IBD algorithm under estimates the FWHM of the PSF while the naïve MLEM blind approach over estimates it. If we increase the number of iterations of the MLEM blind approach, eventually the PSFs tend to disappear completely favoring the Dirac solution. This was discussed at the beginning of Subsection 2.2 and is also shown in Chapter 5. In literature, denoising was applied as a prior step as a form of regularization on the object. It was mentioned that this approach is not recommended because there is no known relationship that models the resulting data to the original observation. In our experiments, we verified that for the same number of iterations of the naïve MLEM BD, denoising before BD over estimates the FWHM of the PSF than the regular approach (*cf.* Fig. 4.8).

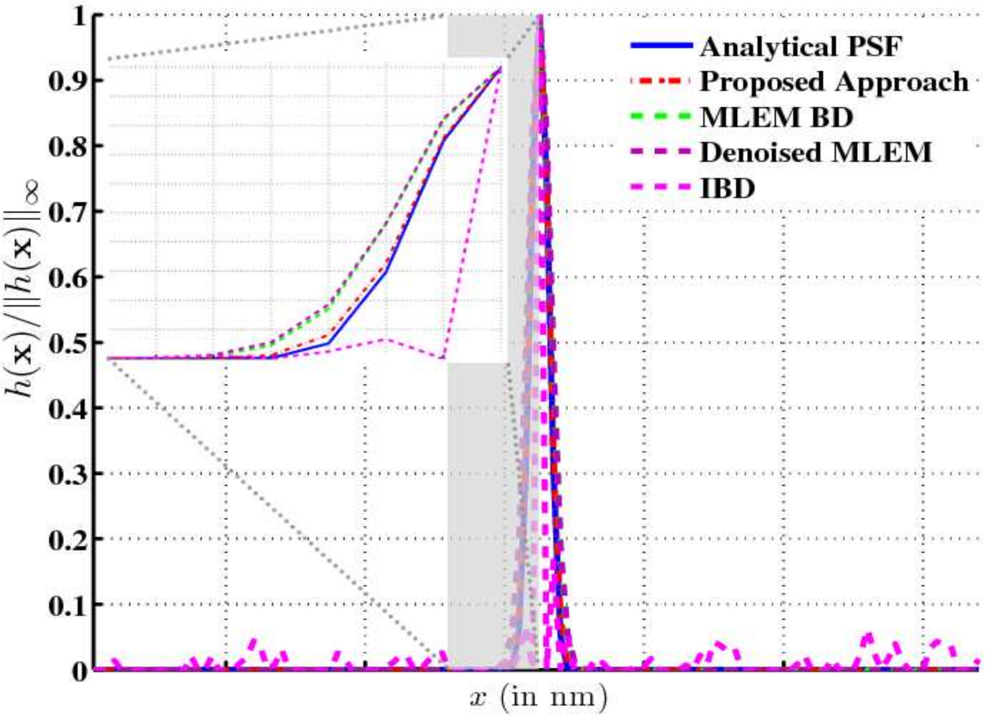
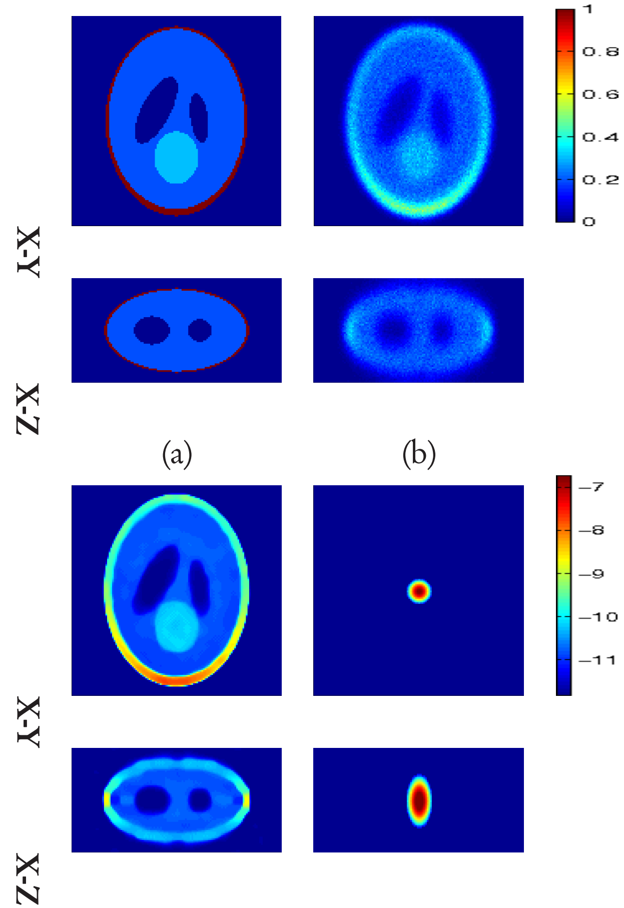


Figure 4.8: ℓ_∞ normalized estimated PSFs obtained from the blind deconvolution algorithms compared with the given analytical model (©Ariana-INRIA/I3S).

In Fig. 4.9(c) and (d), we demonstrate the deconvolution algorithm on simulation with low photon counts (with $\gamma = 10$). For this case, λ_o was estimated to be 0.00451. Fig. 4.10 shows the reduction in the cost function with iterations of the GD algorithm and the approach of the estimated lateral spread parameter σ_r to the stable value, given the estimate of the object. The quality of the restoration can be assessed by using the i-divergence in Eq. (1.18) to compare the restoration with the original synthetic object. When the stopping criterion ϵ was reached,



(a) (b)

(c) (d)

Figure 4.9: (a) 3D phantom object shown with false coloring, (b) observed image blurred by the analytical PSF model (Stokseth [1969]) and Poisson noise ($\gamma = 10$, $\mathrm{IDIV}(i,o) = 5.5544$), (c) image restored after MLEM+TV deconvolution using the estimated PSF ($\mathrm{IDIV}(\hat{o},o) = 1.4334$), (d) estimated diffraction-limited PSF. The original object, observed image and the restoration result use the same linear scaling of intensities while the PSF is on a logarithmic scale (ⓒ Ariana-INRIA/I3S).

the final $\mathrm{IDIV}(o|\hat{o})$ between o in Fig. 4.9(a) and $\hat{o}$ in (c) is 1.4334. Fig. 4.11(a) compares the estimated 3D PSF with the analytically modeled PSF in Eq. (3.52) and the best 3D Gaussian fit (in the least-squared sense) for the analytical model. The PSFs are shown along one direction of an off-central lateral plane, and the highlighted section of the plot can be viewed in the inset. The maximum of the residual error between the estimate and the true PSF is displayed on a logarithmic contrast stretch in Fig. 4.11(b). As is evident from the log of the residue, the Gaussian model does not capture the side lobes in the analytical model, and yet

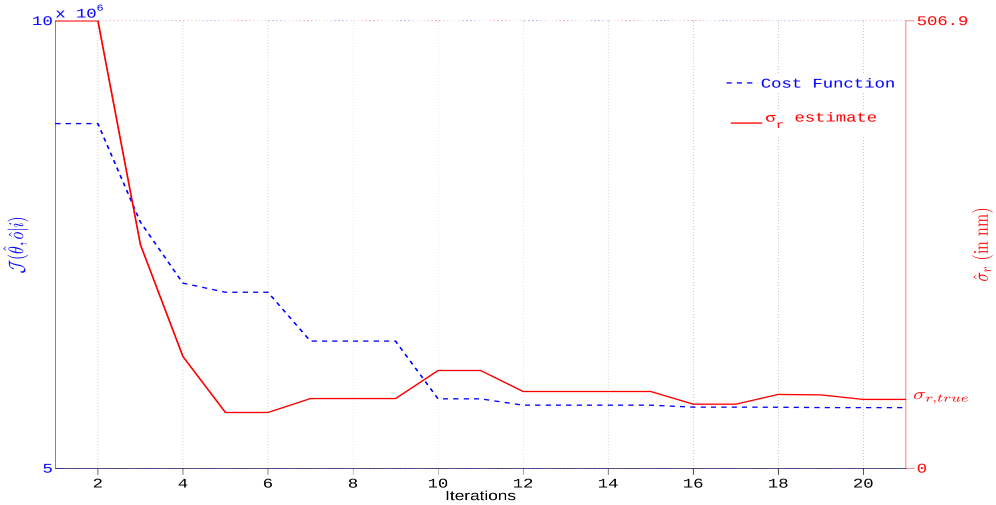


Figure 4.10: Progress of the AM algorithm iteration showing convergence of the cost function and lateral parameter by the GD method (when the original object is known). The Y axis is left-scaled for the cost function $\mathscr{J}(\hat{\omega}_b, \hat{o}|i)$ and right-scaled for the PSF parameter respectively (©Ariana-INRIA/I3S).

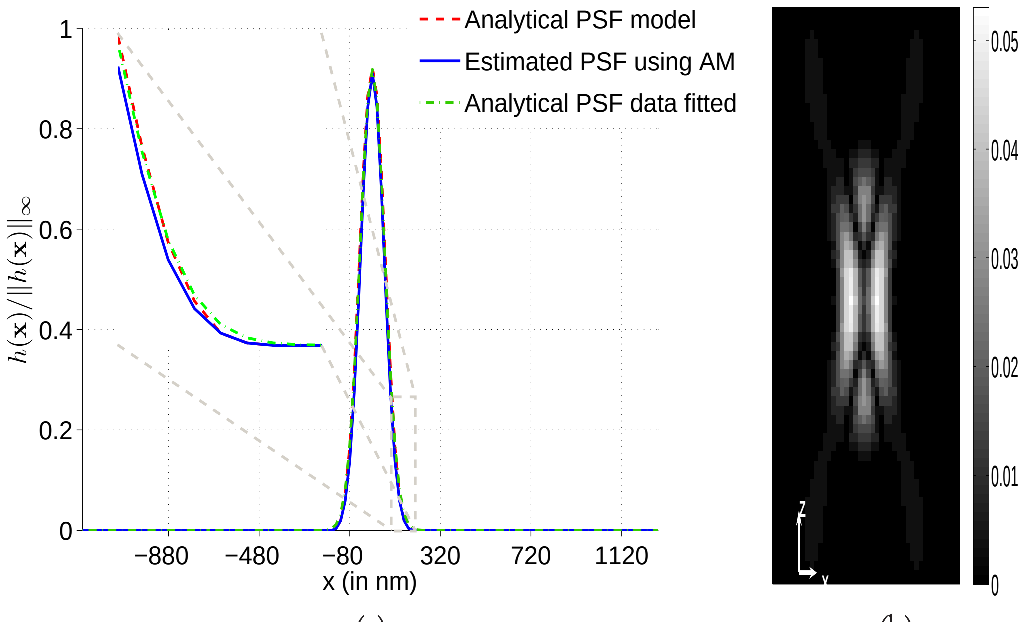
the RSE was found to be $< 0.07\%$. We noticed that the contour of the edges in the deconvolved results are not restored well. The reason for this is that the object we chose initially for simulation is not a band-limited object while the deconvolution algorithm tries to approximate it as a band-limited object. This conflict makes it appear like there is an artifact in the deconvolved result when in reality this result is expected. We will discuss more on this subject in the next chapter.

4.3.3 Experiments on Fluorescent Microspheres

Validation of the algorithm is very important because it would be hard to say if the observed structures are artifacts from restoration or actual biological structures. However, the results on real data are difficult to be validated unless a higher resolution image of the same sample is available or the object that we are scanning is known *a priori*. Hence, we dedicate this subsection to test our algorithm on observed images of a fluorescent microsphere.

Imaging Setup

The Zeiss LSM 510 confocal microscope is mounted on a motorized inverted stand (Zeiss Axiovert 200M) and is equipped with an *ArKr* excitation laser of



(a) (b)

Figure 4.11: (a) The analytically modeled (dash), estimated (continuous) and the best 3D Gaussian fit PSFs (dash-dot) are displayed for one direction, and the inset shows a section of the plot (highlighted) for better visual inspection; (b) X-Z projection of the residual between the estimated and analytically modeled PSF is (ℓ^2 of residual $< 0.07\%$) displayed on a logarithm contrast stretch (ⓒAriana-INRIA/I3S).

wavelength of 488nm. The band pass (BP) filter transmits emitted light within the band 505 − 550nm.

The acquisition software is Zeiss LSM 510 Meta®, which stores the images as 2D stacks. The LSM file format used is a set of 2D tiff images which could be read using the standard image processing libraries. The header of this file contains the acquisition information. As the image acquisition is done by raster scanning, it is possible to have one or several passes for each line in the image and then average the scans. Although this reduces the noise, but it can also destroy the sample due to photobleaching. Line averaged observation also do not satisfy the hypothesis we chose for the observation. Thus, no line averaging was done on any of the real data chosen for testing the algorithm.

Fluorescent Microspheres

These microspheres, manufactured by Invitrogen™, have a specified nominal diameter of 15μm. Molecular Probes® uses its proprietary fluorescent dye technology to produce these polystyrene spheres coated with a fluorescent dye. The fluorescence of the sample can be measured by exciting the dye at a specific wavelength and recording the emission light at their respective maximum emission wavelengths. Since these dyes have a penetration depth between 500 − 700nm, and on emission, they resemble a spherical shell. These images are very handy because the object viewed has a definite geometry and size. In Fig. 4.13(a), we show

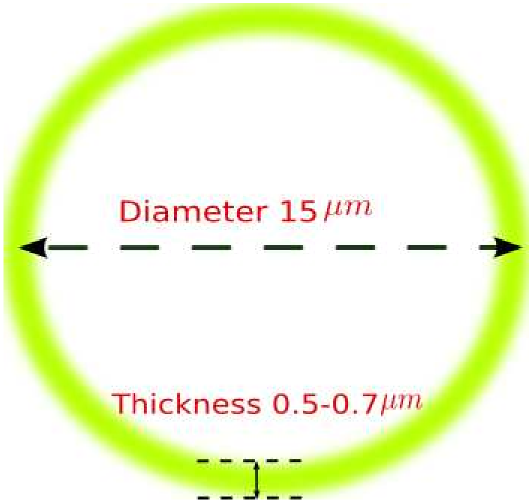
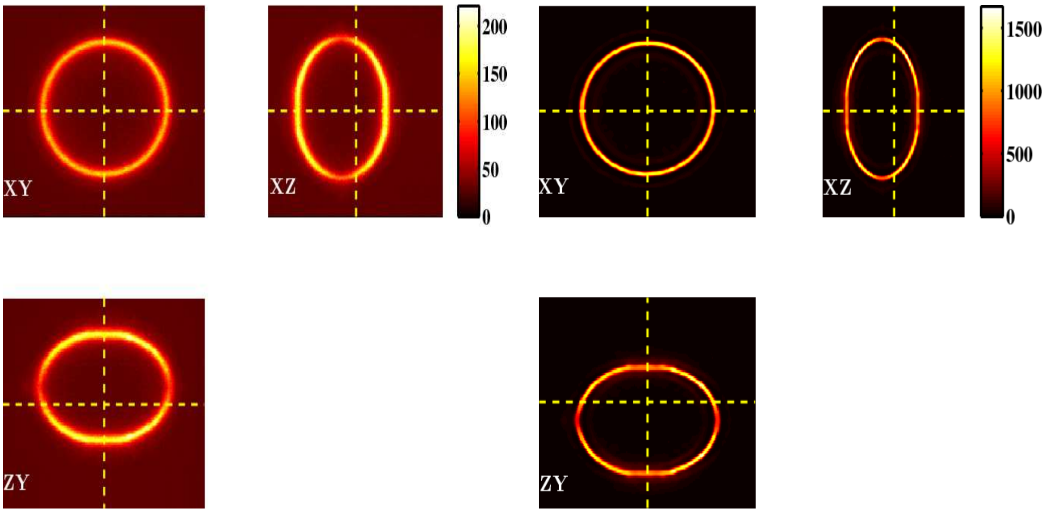


Figure 4.12: Illustrative diagram of a fluorescent microsphere used for the experiment. The manufacturer specified diameter of the sphere is 15μm and it is surface coated with a fluorescent dye having a penetration depth between 500 − 700nm (©Ariana-INRIA/I3S).

the raw data that is Nyquist sampled at 89.26nm along the lateral and 232.75nm along the optic axis. The excitation peak wavelength, in this case, is 488nm and emission is at 520nm. The backprojected pinhole size is 1AU and no line averaging was done during the acquisition process. The total size of the image stack is $256 \times 256 \times 130$ voxels or $22.85 \times 22.85 \times 30.03\mu$m. The objective lens used is an oil immersion '*Plan Apochromat*' with NA 1.4 and 63× magnification. Although the geometry of this test sample is nice, as the RI of the sphere is in general not the same as that of the embedding media, they also degrade the PSF. In particular, when viewing the central plane, those rays approaching the focus from the direction of the center of the sphere will be highly aberrated and will produce a radially asymmetrical PSF. Thus, images of the far side of the sphere will be different than those of the near side. In order to remove these aberrations, only one half of the volume stack was used. The whole volume was then regenerated by mirroring only the upper half of the volume about the central axial plane.



(a) (b)

Figure 4.13: Blind deconvolution of the observed microsphere images. (a) Raw observed images with voxel sizes $\Delta_{xy} = 89.26$nm and $\Delta_z = 232.75$nm, (b) the processed data using proposed approach with $\hat{\lambda}_o = 0.0046$ and IDIV$(\hat{o}, i) = 10.2$ (©Ariana-INRIA/I3S).

We define the thickness of the shell by using the FWHM of the line profile peak. For the observed images, the FWHM was measured as 930nm along the central radial plane. This is much larger than the manufacture specified range. In Dey, et al. [2004], these images were processed with a non-blind deconvolution algorithm, and a numerically calculated PSF was used. It was mentioned that when these images were processed using a deconvolution algorithm without regularization, the FWHM was 260nm, while with TV regularization, the thickness was much closer at 400nm.

Microsphere Deconvolution Results

In Fig. 4.13(b), we show the deconvolution results obtained using our approach. For this case the parameter $\hat{\lambda}_o$ is estimated as 0.0046. The estimated background $\hat{b}$ is 30 IU, and the AM algorithm is continued until the convergence criterion is

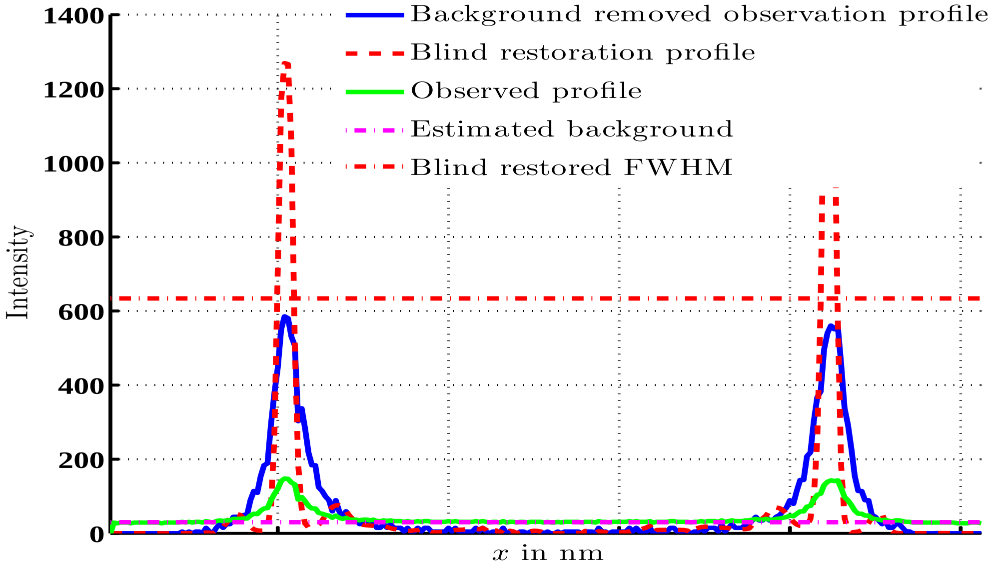


Figure 4.14: Intensity profile of the central axial plane for: the observation, background subtracted and normalized observation, and the blind restored volume (©Ariana-INRIA/I3S).

satisfied. We plot the intensity profiles for the observation and the restoration in Fig. 4.14. The horizontal line in the figure which is shown to bisect the restoration intensity profile is the line at the half intensity maximum. When the width between the intersection points was measured, it was found to be 535.58nm. In fact, this value lies between the manufactured specified range of the penetration depth $500-700$nm. We thus demonstrate the effectiveness of restoring known objects without causing any artifacts and also our algorithm shows at least twofold improvement in the resolution.

4.3.4 Experiments on Real Data

Sample Description

There were two specimens that were chosen for testing the BD algorithm. The first is an embryo of the *Drosophila Melanogaster* (*cf.* Fig. 4.15(a)). It is mounted and tagged with the GFP. This preparation is used for studying the sealing of the epithelial sheets (dorsal closure) midway during the embryogenesis. The objective lens is a Plan-Neofluar with 40× magnification having a NA of 1.3 and immersed in oil (Immersol$^{\text{TM}}$ 518F, Zeiss, RI $n_i = 1.518$). The pinhole size is

67μm. The images [Institute of Signaling, Development Biology and Cancer, Nice UMR6543/CNRS/UNS] were from the sample just below the coverslip, and acquired with a XY pixel size of 50nm and a Z step size of 170nm. The size of the volume imaged is $25.59 \times 25.59 \times 2.55\mu$m.

The second set of images [National Institute for Agricultural Research (INRA), Sophia-Antipolis] are the root apex of the plant *Arabidopsis Thaliana* immersed in water (*cf.* Fig. 4.16). The dissected roots of the plant were directly put on a microscope slide in approximately 100μl of water and this was then gently covered with a coverslip. This simple set up works very well when the image acquisition recording times are not too long (about 30 mins). The microscope specifications are the same as that used for acquiring the first data set but the objective is a C-Apochromat water immersion lens with 63× magnification, 1.2 NA. The lateral pixel dimensions are 113nm and the Z step is 438nm for a pinhole size of 110μm. This preparation was used to study *Nematode* infection at the center of the root in the vascular tissue.

Deconvolution Results

A rendered sub-volume of observed and restored images for the *Drosophila Melanogaster* is shown in Fig. 4.15. The observation contained a high level of noise, and the optimum value of the regularization parameter λ_o is estimated to be 0.0047. The deconvolution algorithm was stopped when the difference between subsequent estimates was lower than $\epsilon = 0.002$. The AM algorithm converged after 40 iterations of the joint MLEMTV and the GD algorithm. The PSF parameters were initialized to 300nm and 600nm for the lateral and the axial case respectively, and the GD algorithm estimated them to be 257.9 and 477.9nm (Pankajakshan, et al. [2008a, 2009b]). These are larger (by about 16% and 14.5% for the lateral and the axial case respectively) than their corresponding theoretically calculated values (Zhang, et al. [2007]). These results are also fully in agreement with an independent experimental study performed by de Moraes Marim, et al. [2008] with sub-resolution beads, which indicated a deviation of theoretical aberration-free PSF models from empirically determined PSFs.

Fig. 4.17(a) shows a rendered sub-volume (as indicated in Fig. 4.16) of the observed root apex and the corresponding restored result is shown in Fig. 4.17(b). It

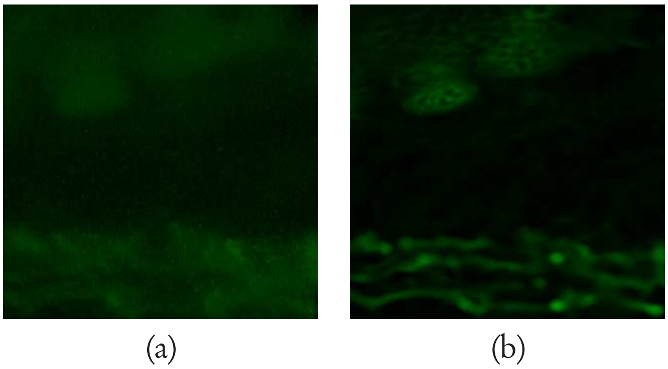

(a) (b)

Figure 4.15: (a) Rendered sub-volume of the original specimen (©Institute of Signaling, Developmental Biology & Cancer UMR6543/I3S/UNS), and (b) restored image (©Ariana-INRIA/I3S). The intensity is scaled to $[0, 130]$ for display.

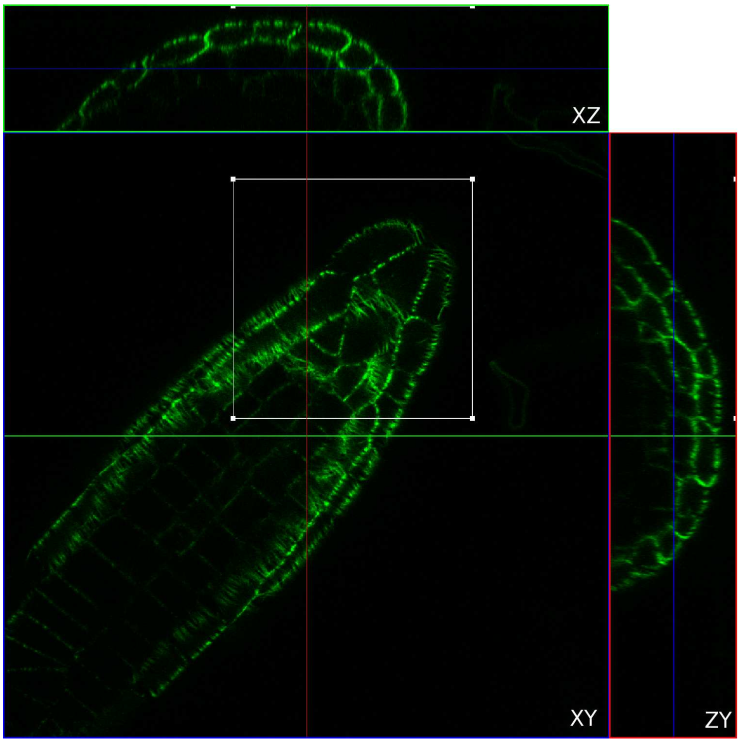


Figure 4.16: Observed root apex of an *Arabidopsis Thaliana* with a volume 146.448μm $\times$ 146.448μm $\times$ 30.222μm (©INRA). The sub-volume chosen for restoration is emphasized.

is evident from these results that the *microtubules* [as identified by their specific binding proteins (*cf.* Marc, et al. [1998]), microtubules binding domain (MBD)] are much easily discerned in the restoration than in the original data.

In Subsection 1.3.3, with the help of observed images under different pin-

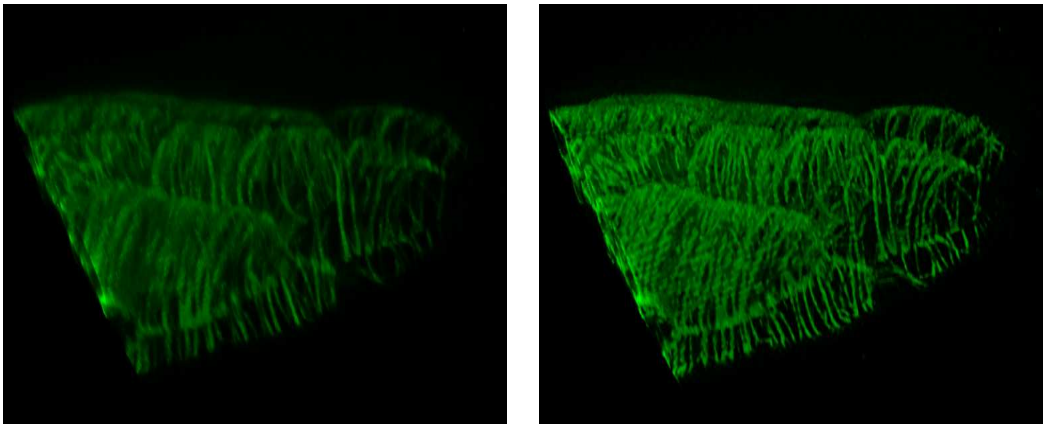
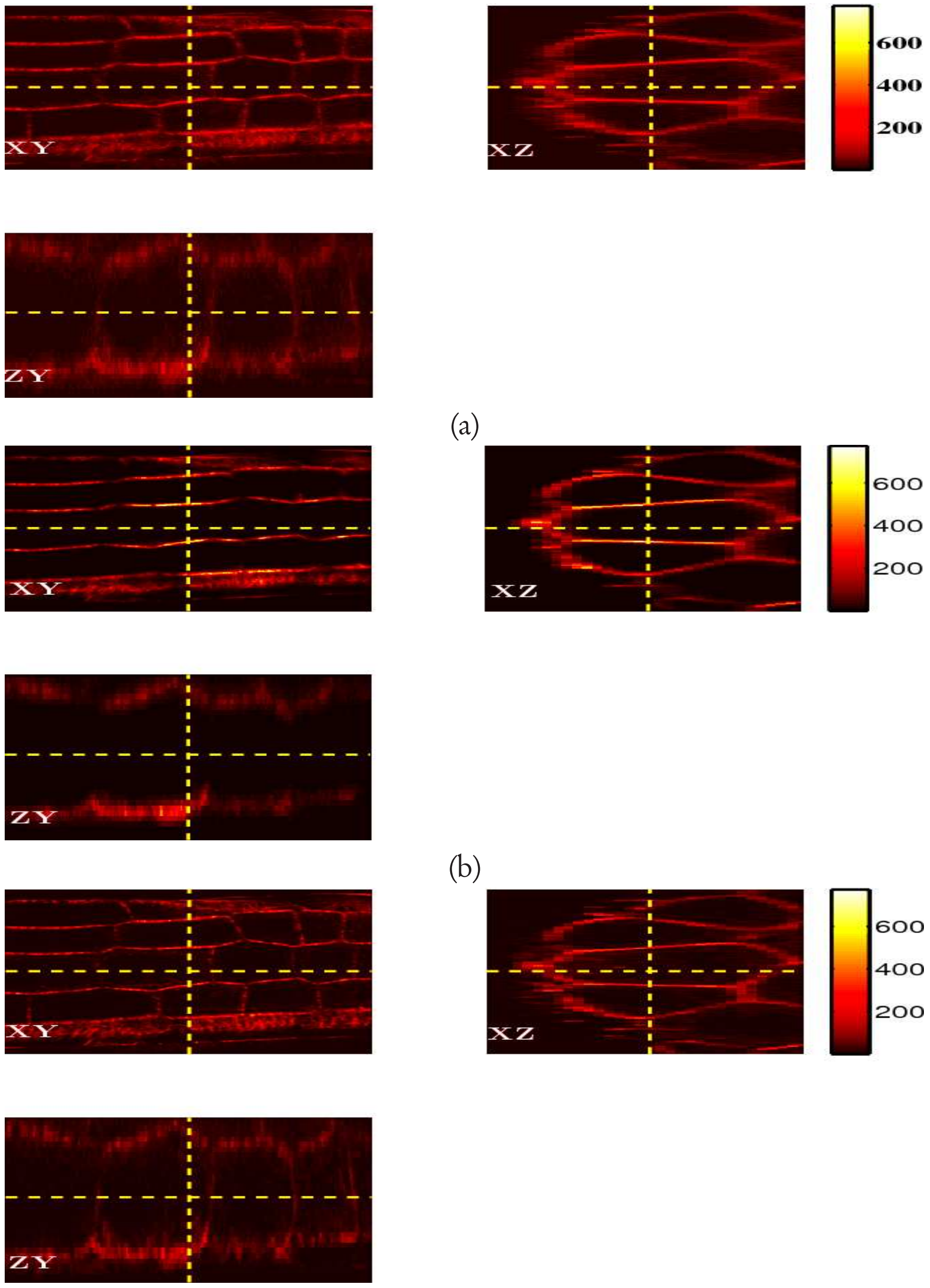

(a) (b)

Figure 4.17: Rendered subvolume of the (a) observed image slices in Fig. 4.16 (ⓒINRA) and (b) restored image slices (ⓒAriana-INRIA/I3S). $\lambda_o = 0.002$ and $\epsilon = 0.0001$.

hole settings, we highlighted the trade-off biologists have to make between pinhole size, resolution, noise and speed of scanning. We mentioned how we could consider the observation obtained from the 1AU pinhole setting as high resolution (albeit noisy), and the image's z resolution degrades with increase in the pinhole size but with SNR improving. We consider the microscope to the WFM under the limiting case of a completely open pinhole. We tested our algorithm on the observed images of the same plant specimen, but now with a pinhole size of 2AU. The size of the chosen volume in Fig. 4.18(a) was $256 \times 256 \times 32$ voxels, with the lateral and axial sampling sizes of 285.64nm and 845.62nm. The theoretically calculated PSF parameters from Subsection 4.2.2 were $(\hat{\sigma}_{r,\mathrm{Th}}, \hat{\sigma}_{z,\mathrm{Th}}) = (498.32, 970.922)$ nm while our estimated parameters are $(\hat{\sigma}_{r,\mathrm{BID}}, \hat{\sigma}_{z,\mathrm{BID}}) = (802.30, 2851.90)$ nm. The parameters obtained by a 3D anisotropic Gaussian fit to the PSF, $\hat{h}_{\mathrm{MLEM}}$, estimated by using naïve MLEM (from Eq. (2.35)) is $(\hat{\sigma}_{r,\mathrm{MLEM}}, \hat{\sigma}_{z,\mathrm{MLEM}}) = (147.52, 366.60)$nm. We found that the results obtained in Fig. 4.18(c) is comparable with the resolution obtained using the 1AU pinhole in Fig. 1.7(a). We see that the advantage of using the deconvolution algorithm is that we can do fast scanning of large samples with lesser noise and sometimes much better resolution. We show the naïve MLEM BD algorithm results as a reference for comparison in Fig. 4.18(b) because, we noticed that cell walls along specific directions were lost completely even though these were present in the original observation. Although, there is some loss in contrast



(a)

(b)

(c)

Figure 4.18: (a) Observed section of a volume with 2AU pinhole, (b) blind naïve MLEM approach (*cf.* Holmes [1992]; Biggs & Andrews [1997]), (c) proposed approach (©Ariana-INRIA/I3S). As the samples are thin, in this figure, the z sections were enlarged in size for ease of viewing.

with our proposed approach, the structure of the plant cell walls are maintained and enhanced.

4.4 Conclusion

In this chapter, we proposed and validated an alternate minimization (AM) algorithm for the joint estimation of the microscope PSF and the specimen source distribution for a CLSM. We choose the MLEM algorithm for the deconvolution process as it is best suited for the Poisson data, and TV as the regularization model. As inference of the PSF directly from the observation is difficult, we recommend using a separable 3D Gaussian model for the diffraction-limited confocal PSF. This is a better *a priori* model than introducing of constraints on the OTF space because the imaging parameters are not necessary. The PSF approximation that is given in this chapter is currently relevant to imaging thin samples. However, they could also be extended to encompass any PSF that can be decomposed in a similar manner. We experimented on simulated, microsphere images and real data. For the simulated case, the method gives very good deconvolution results and a PSF estimation close to the true value. There were some boundary artifacts but this was caused due to the estimation of non-bandlimited synthetic object. A point that we would like to drive home from these results is that it is imperative to use a regularization on the object that we would like to restore. In the absence of regularization, as in the case of the naïve MLEM BD, the algorithm does converge but not to the trivial solution. Our proposed algorithm was validated on a set of microsphere images with known fluorescence penetration depth and object geometry. We were able to estimate the fluorescence penetration thickness after BD, and it was found to lie within the levels specified by the manufacturer. BD results from low resolution data obtained with 2AU pinhole was comparable to the high resolution data obtained by closing the pinhole to 1AU. Biologists who wish to acquire images faster, to prevent photobleaching, may acquire lower resolution images and use the BD to obtain higher resolution images with minimum noise. The TV regularization can sometimes causes a loss in the contrast of the restored images. In order to overcome this difficulty, recently, an extension to this approach was proposed by Brune, et al. [2009]. Their method enhances the contrast by using the concept of inverse scale spaces and Bregman [1965] iterations.

CHAPTER 5

Pupil Phase Retrieval by Parameter Estimation

> *"There are two kinds of light-the glow that illumines, and the glare that obscures."*
>
> *-James Grover Thurber (American author and cartoonist)*

The images of fluorescent microspheres are often used for deconvolving an observed specimen volume. However, the qualities of such a derived PSF are dependent on the lens type, alignment, correct coverslip thickness, immersion medium, proximity of the specimen to the coverslip, and uniformity of the illumination (Shaw [2006]). If the size of the microsphere is too small (sub-resolution), then the observation is plagued with noise problems as the microsphere is only weakly fluorescent. Lai, et al. [2005] noticed that an iterative deconvolution algorithm is very sensitive to the amount of randomness in the experimental PSF. It can also be shown that the recovered objects are significantly improved by using a smoothed experimental PSF rather than using the experimental PSF directly. Similarly, for the algorithm tested on simulated objects, it was found that the MSE$(o,\hat{o})$ is significantly smaller for a filtered experimental PSF than using the non-filtered one. Non sub-resolution spheres cannot be directly used as they are too large to represent an ideal point source.
The motivation for this chapter is as follows:

- knowledge of the PSF of an imaging system helps in recovering the original intensity distribution of an imaged object by deconvolution,
- experimentally obtained PSFs by imaging fluorescent spheres are low in contrast and highly noisy,

- depending on the microsphere size, the derived PSFs are larger in size than the true PSF, and
- these derived PSFs cannot be reused if the experimental conditions are varied.

Drawing inspiration from Chapter 4 on the subject of blind deconvolution, we propose an approach based on Bayesian inference to recover both the imaged object and the PSF from the experimental data. The difference here is that we are more interested in estimating PSFs by taking into account spherical aberrations (SA); but with some knowledge of the object. The observed objects have a specific geometry and manufacture specified parameters. An extension of this work will involve using this method on the images of biological specimens affected by SA.

5.1 Introduction

Theoretically calculated PSFs often lack experimental or microscope specific signatures, while the empirically obtained images of microspheres have larger FWHM (than theoretically predicted), if the microspheres are large, or (and) too noisy if microspheres are too small (*cf.* Fig. 5.1 for example). As explained in Section 3.3, the non-aberrated PSF model in Eq. (3.27) is valid only when imaging thin samples (NFP $< 20-100\mu$m). When imaging into deeper sections within a biological sample, in confocal or nonlinear microscopy, sample induced aberrations are introduced. There is a dramatic reduction in both the signal level and the resolution. While most systems are built to be diffraction-limited, it is not possible to ensure that they maintain this performance throughout. The observed image quality is influenced by both the diffraction effect and SA. It was shown by Pawley [2006] that when focusing into samples immersed in water by using an oil immersion objective, the signal intensity level falls to 40% at 5μm and to 10% at 15μm beneath the coverslip (Booth, et al. [1998]). This is because, the RI mismatch between the specimen and the immersion medium of the objective lens becomes significant (Egner, et al. [1998]). This mismatch produces an additional path difference in the emerging light wavefront. Hence, we use the modified PSF expression in Eq. (3.27) to also include dependence on the depth as in Eq. (3.45).

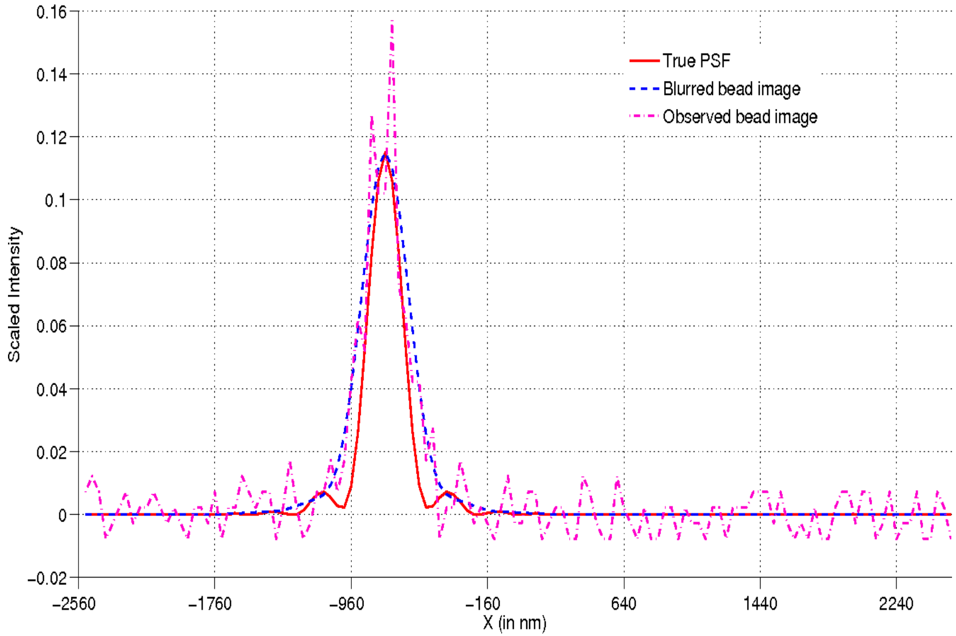


Figure 5.1: Axial intensity profiles of a calculated confocal PSF, a blurred microsphere and the observed microsphere. The intensities are scaled and the peaks matched for visual comparison (ⓒAriana-INRIA/I3S).

This chapter is organized as follows. In Section 5.2, we modify the previously proposed joint maximum likelihood (JML) and joint maximum *a posteriori* (JMAP) to retrieve the pupil phase. The algorithm presented here retrieves the pupil phase with a Bayesian viewpoint, and it draws its inspiration from the works Hanser, et al. [2001, 2003] for WFM. We will provide some existing literature on the subject of BD, and we shall later see why these methods cannot be applied for solving the above problem. To test our hypothesis, in Section 5.3, we simulated a bandlimited object and generated the observation for experimentation (*cf.* Appendix A). The iterative algorithm needs a good initialization and this is handled in Subsection 5.3.2. We also show that this new model, when used for deconvolution, provides a much more realistic reconstruction than the conventional model. In Subsection 5.3.3, we will see how the existing methods of deconvolution fail in the presence of aberrations. The role of the phase of the emerging refracted wavefront in restoring the observed objects will also become much more evident from the examples. Finally, we discuss the possibility of extending this work to retrieve the PSF directly from observed specimen data and conclude in Section 5.4.

5.2 Phase Retrieval with a Bayesian Viewpoint

In literature, prior work related to this subject is mostly on wavefront reconstruction using adaptive optics (AO) (*cf.* Hiraoka, et al. [1990]; Kam, et al. [1997]; Sherman, et al. [2002]; Booth, et al. [2002]; Kam, et al. [2007]). These AO methods are based on the idea of phase aberration compensation by physically adding a special conjugate element in the optical system. A review of the recent trends in this field was done by Booth [2007]. Such a compensation could also be computationally achieved as shown by Kam, et al. [2001]. In Juskaitis [2006], the amplitude and phase of the pupil function was measured by using a fiber-optic interferometer. If either of this information is available, deconvolution could be achieved for the samples that are imaged by using similar experimental settings. In the following section we will discuss an approach of phase retrieval without the use of such a complex setup and also extend it to the case of directly restoring them from biological specimen images.

5.2.1 Joint Maximum Likelihood Approach

In Pankajakshan, et al. [2009a], we used the MLEM for removing the effect of the microsphere size and later estimated the pupil phase parameters by minimizing a cost function. The MLEM algorithm was discussed in Subsection 2.1.2 for the case of direct object estimation from the observation data. The same approach was used but the roles of the object and the PSF were reversed. In this chapter, however, we will use a joint minimization technique that treats the problem of size compensation inherently. Since the object that we are interested in restoring is a microsphere, its geometry and manufacture specified diameter is known *a priori*. We denote by $\omega_o \in \Theta$ the unknown parameter(s) of the object, and by $\omega_h \in \Theta$ those of the pupil phase. The knowledge of these parameters allow us to completely define the object and the PSF. With the parametrization of the object and the PSF, the ML approach in Subsection 2.2.3 cannot be used in its direct form. We thus rewrite the likelihood in terms of these parameters as

$$\Pr(i|\omega_o,\omega_h)=\prod_{\mathbf{x}\in\Omega_s}\frac{\left([(h(\omega_h)*o(\omega_o))+b](\mathbf{x})\right)^{i(\mathbf{x})}\exp\left(-[(h(\omega_h)*o(\omega_o))+b](\mathbf{x})\right)}{i(\mathbf{x})!}. \tag{5.1}$$

By applying the -log operator to Eq. (5.1), we get the following energy functional:

$$\mathcal{J}_{\text{obs}}(i|\omega_o,\omega_h)=\sum_{\mathbf{x}\in\Omega_s}\left[(h(\omega_h)*o(\omega_o))+b\right](\mathbf{x})-$$
$$\sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x})\log\left[(h(\omega_h)*o(\omega_o))+b\right](\mathbf{x})+\sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x})!\,. \quad (5.2)$$

Maximizing the likelihood in Eq. (5.1) with respect to the parameters ω_o and ω_h and minimizing the cost function, $\mathcal{J}(i|\omega_o,\omega_h)$, is equivalent. Thus,

$$(\hat{\omega}_o,\hat{\omega}_h)=\underset{(\omega_o,\omega_h)\in\Theta}{\operatorname{argmax}}\ \{\Pr(i|\omega_o,\omega_h)\} \quad (5.3)$$
$$=\underset{(\omega_o,\omega_h)\in\Theta}{\operatorname{argmin}}\ \{\mathcal{J}_{\text{obs}}(i|\omega_o,\omega_h)\}\,. \quad (5.4)$$

The term $i(\mathbf{x})!$, in Eq. (5.1), is independent of the parameters to be estimated and we drop it from the maximization process.

5.2.2 Joint Maximum *A Posteriori* Approach

It was shown by Dainty & Fiddy [1984] that when the observed intensity data are available, unique reconstruction of the phase of the pupil is possible provided there is *a priori* information on the phase or the object support. As the solution space for the parameters is large, we add some realistic priors on the parameters to restrict the possible outcomes. The Bayes' theorem can be used to estimate the unknown parameters ω_h and ω_o from the observation i. In this we treat the observation, object parameters and the PSF parameters as probability-frequency functions. Thus, using the Bayes' inference, the conditional posterior probability is

$$\Pr(\omega_o,\omega_h|i)=\frac{\Pr(i|\omega_o,\omega_h)\Pr(\omega_o)\Pr(\omega_h)}{\Pr(i)}, \quad (5.5)$$

where $\Pr(i)$ is the prior on the observation. In the above expression, we assumed that the object and the PSF parameters are independent of each other. So each of them could be independently treated with hyperprior distributions $\Pr(\omega_o)$ and $\Pr(\omega_h)$. The estimates for ω_o and ω_h can be obtained by the joint maximum *a* *posteriori* (JMAP) or minimizing the cologarithm of the posterior as

$$\begin{aligned}(\hat{\boldsymbol{\omega}}_o,\hat{\boldsymbol{\omega}}_h) &= \underset{(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h)\in\Theta}{\operatorname{argmin}} \{\Pr(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h|i)\}\\ &= \underset{(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h)\in\Theta}{\operatorname{argmin}} \{-\log[\Pr(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h|i)]\}\,. \end{aligned} \tag{5.6}$$

The energy function in Eq. (5.4) can now be rewritten as sum of three individual energy terms: the first is the data-fidelity term obtained from the observation model Eq. (5.1), the second and the third energy terms, $\mathcal{J}_{\mathrm{hp}}(\boldsymbol{\omega}_o)$ and $\mathcal{J}_{\mathrm{hp}}(\boldsymbol{\omega}_h)$, are generated from the hyperpriors, $\Pr(\boldsymbol{\omega}_o)$ and $\Pr(\boldsymbol{\omega}_h)$.

$$(\hat{\boldsymbol{\omega}}_o,\hat{\boldsymbol{\omega}}_h) = \underset{(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h)\in\Theta}{\operatorname{argmin}} \{\mathcal{J}(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h|i)\} \tag{5.7}$$

$$= \underset{(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h)\in\Theta}{\operatorname{argmin}} \{\mathcal{J}_{\mathrm{obs}}(i|\boldsymbol{\omega}_o,\boldsymbol{\omega}_h)+\mathcal{J}_{\mathrm{hp}}(\boldsymbol{\omega}_o)+\mathcal{J}_{\mathrm{hp}}(\boldsymbol{\omega}_h)\}\,. \tag{5.8}$$

It is easy to note that Eq. (5.8) reduces to Eq. (5.4) when flat hyperpriors are used. On the other hand, if we assign a Gamma prior distribution to each of the parameters, then

$$\Pr(\omega_o)\sim\omega_o^{a_o-1}\exp(-b_o\omega_o), \tag{5.9}$$

and

$$\Pr(\omega_h)\sim\omega_h^{a_h-1}\exp(-b_h\omega_h), \tag{5.10}$$

where $(a_{(\cdot)},b_{(\cdot)})$ are the parameters of the Gamma prior. The generalized energy function in Eq. (5.8) is written by using the expressions in Eqs. (5.2), (5.9) and (5.10) as

$$\begin{aligned}\mathcal{J}(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h|i) = &\sum_{\mathbf{x}\in\Omega_s}[(h(\boldsymbol{\omega}_h)*o(\boldsymbol{\omega}_o))+b](\mathbf{x})\\ &-\sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x})\log[(h(\boldsymbol{\omega}_h)*o(\boldsymbol{\omega}_o))+b](\mathbf{x})+\\ &\sum_{l=1}^{N_o}(a_{o,l}-1)\log(\omega_{o,l})-b_{o,l}\omega_{o,l}+\\ &\sum_{m=1}^{N_h}(a_{h,m}-1)\log(\omega_{h,m})-b_{o,m}\omega_{o,m}\end{aligned} \tag{5.11}$$

where $a_{o,l}$ and $b_{o,l}$ are the parameters of the hyperprior for the l^{th} parameter $\omega_{o,l}$ of the object. Similarly, $a_{h,m}$ and $b_{h,m}$ are the parameters of the hyperprior for the m^{th} parameter $\omega_{h,m}$ of the PSF. In the above equation, we assumed these parameters to be known *a priori*. This need not necessarily be the case.

In this MAP approach, the background fluorescence term b is determined from a single "specimen-independent" slice as $\hat{b}$. It is only estimated once for the whole volume, and is then assumed to be known during the estimation of the object $\boldsymbol{\omega}_o$ and the PSF parameters $\boldsymbol{\omega}_h$. This estimation procedure is as described in Subsection 1.3.2.

Object and Point-Spread Function Parameters Estimation

An observed source point appears shifted in depth from its true axial position due to SA. If we assume that the sphere placed at a relative position (x_o, y_o, z_o) in a given volume, then the observation will have its centroid in the image plane position (x_i, y_i, z_i), with $x_i \approx x_o$, $y_i \approx y_o$ and $z_i \neq z_o$. In the spatial domain, the object can be written as

$$o(x,y,z) = \begin{cases} 1, & \forall \left((x-x_o)^2+(y-y_o)^2+(z-z_o)^2\right)^{\frac{1}{2}} \leq R \\ 0, & \text{otherwise}, \end{cases} \tag{5.12}$$

where R is the radius of the sphere specified by the manufacturer. In the above case, we assumed the true intensity to be unity and uniform within the sphere. However, the true intensity need not necessarily be unity. We recall that due to photon loss, the true uniform intensity of the observed sphere is unknown. Thus, the two object parameters to be estimated are the intensity and the relative position. If we assume that the apodization function A in Eq. (3.28) is uniform (as in Subsection 3.4.3), this term can be absorbed into the intensity of the sphere in Eq. (5.11). So this combined term to be estimated will be represented by a scalar s, where $s \in \mathbb{R}$. In the case when the true intensity is unity, the apodization function A will be a uniformly constant value of s. The estimation of this entity will be discussed separately in Subsection 5.2.2. So for the object, the only parameter that needs to be estimated is the relative position z_o, and hence $\boldsymbol{\omega}_o = \{z_o\}$.

In the presence of SA, for a calibrated CLSM, it was mentioned in Subsection 3.4.2 that the variation of the PSF is essentially a factor of the following three parameters: RI of the objective immersion medium n_i, index of the specimen or mounting medium n_s, and depth of the specimen under the cover slip d. Thus, the PSF parameters to be estimated are $\boldsymbol{\omega}_h = \{d, n_i, n_s\}$. This was also shown to be the case in the PSF phase expressions derived in Eq. (3.45), and in the nonlinear approximation of Eq. (3.49).

Summarizing, the above discussion, the energy function in Eq. (5.11) can be rewritten as

$$\begin{aligned}\mathscr{J}(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h,s|i) = \sum_{\mathbf{x}\in\Omega_s}[s(h(\boldsymbol{\omega}_h)*o(\boldsymbol{\omega}_o))+\hat{b}](\mathbf{x}) - \\ \sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x})\log[s(h(\boldsymbol{\omega}_h)*o(\boldsymbol{\omega}_o))+\hat{b}](\mathbf{x}) + \\ \sum_{l=1}^{N_o}(\hat{\alpha}_{o,l}-1)\log(\omega_{o,l}) - \hat{\beta}_{o,l}\omega_{o,l} + \\ \sum_{m=1}^{N_h}(\hat{\alpha}_{h,m}-1)\log(\omega_{h,m}) - \hat{\beta}_{o,m}\omega_{o,m}. \quad (5.13)\end{aligned}$$

The object term o in the above expression is a normalized form of the object term o in Eq. (5.11) so that $o : \Omega_s \mapsto [0,1]$.

As there is no closed form expression for calculation of $\boldsymbol{\omega}_o$ and $\boldsymbol{\omega}_h$ from Eq. (5.13), we use a Newton's optimization algorithm (*cf.* Nocedal & Wright [1999]). The Newton's method in optimization is a class of hill-climbing optimization technique that seeks the stationary point of a function, where the gradient of the function is zero. However, the Newton method is based on second-order Taylor's expansion of the penalty variation for a small change of parameters:

$$\begin{aligned}\delta\mathscr{J}_{\boldsymbol{\omega}_o}(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h,s|i) &\overset{\text{def}}{=} \mathscr{J}(\boldsymbol{\omega}_o+\delta\boldsymbol{\omega}_o,\boldsymbol{\omega}_h,s|i) - \mathscr{J}(\boldsymbol{\omega}_o,\boldsymbol{\omega}_h,s|i) \quad (5.14)\\ &= \mathbf{G}(\boldsymbol{\omega}_o)^{\mathrm{T}}\cdot\delta\boldsymbol{\omega}_o + \frac{1}{2}\delta\boldsymbol{\omega}_o^{\mathrm{T}}\mathbf{H}(\boldsymbol{\omega}_o)\delta\boldsymbol{\omega}_o + O(\|\delta\boldsymbol{\omega}_o\|_2^2), \quad (5.15)\end{aligned}$$

where $\mathbf{G}(\boldsymbol{\omega}_o) \overset{\text{def}}{=} \nabla_{\boldsymbol{\omega}_o}\mathscr{J}(\boldsymbol{\omega}_o,\hat{\boldsymbol{\omega}}_h|i)$ and $\mathbf{H}(\boldsymbol{\omega}_o) \overset{\text{def}}{=} \nabla^2_{\boldsymbol{\omega}_o}\mathscr{J}(\boldsymbol{\omega}_o,\hat{\boldsymbol{\omega}}_h|i)$ are, respectively, the partial gradient and Hessian of $\mathscr{J}(\boldsymbol{\omega}_o,\hat{\boldsymbol{\omega}}_h,s|i)$. By ignoring the higher order non-quadratic terms of the Taylor expansion, we can show that the optimal step for the Newton's algorithms is: $-\mathbf{H}(\boldsymbol{\omega}_o)^{-1}\mathbf{G}(\boldsymbol{\omega}_o)$. The iterative algorithm for the object parameter estimation can be written as

$$\hat{\omega}_o^{(n+1)} = \hat{\omega}_o^{(n)} - \tau_o^{(n)} \left(\mathbf{H}(\hat{\omega}_o^{(n)})^{-1} \mathbf{G}(\hat{\omega}_o^{(n)}) \right), \text{ s.t } \hat{\omega}_o \in \Theta. \tag{5.16}$$

We note that the previous estimate of $\hat{\omega}_h^{(n)}$ is used in the above iteration. This estimation is followed by the phase parameter iterative algorithm

$$\hat{\omega}_h^{(n+1)} = \hat{\omega}_h^{(n)} - \tau_h^{(n)} \left(\mathbf{H}(\hat{\omega}_h^{(n)})^{-1} \mathbf{G}(\hat{\omega}_h^{(n)}) \right), \text{ s.t } \hat{\omega}_h \in \Theta. \tag{5.17}$$

For the Newton's method, $\nabla^2_{\omega_o} \mathcal{J}(\omega_o, \omega_h | i)$ is the exact Hessian of the cost function with respect to the object parameter ω_o. The calculation of these two terms will be discussed in the next two subsections. Usually Newton's method is modified to include small steps $\tau_o^{(n)}$ and $\tau_h^{(n)}$ at each iteration. This is often done to ensure that the Wolfe conditions (*cf.* Nocedal & Wright [1999]) are satisfied at each step. The initial steps are chosen as $\tau_o^{(0)} = 1$ and $\tau_h^{(0)} = 1$, and they are usually updated in the following manner:

$$\tau_o^{(n+1)} = \frac{\tau_o^{(n)}}{2}; \tau_h^{(n+1)} = \frac{\tau_h^{(n)}}{2}. \tag{5.18}$$

The Eqs. (5.16) and (5.17) are done alternatively so that every iteration n

$$\mathcal{J}(\hat{\omega}_o^n, \hat{\omega}_h^n | i) < \mathcal{J}(\hat{\omega}_o^{n-1}, \hat{\omega}_h^{n-1} | i). \tag{5.19}$$

The geometric interpretation of Newton's method is that at every iteration, $\mathcal{J}(\hat{\omega}_o, \hat{\omega}_h | i)$ is approximated by a quadratic function around $\hat{\omega}_h$ or $\hat{\omega}_o$.

Gradient Calculation The general form of the energy function's partial gradient as a function of the parameters ω_o and ω_h can be written as

$$\begin{aligned} \nabla_{\omega_o} \mathcal{J}(\omega_o, \omega_h | i) = \sum_{\mathbf{x} \in \Omega_s} h(\mathbf{x}; \omega_h) * \frac{\partial}{\partial \omega_o} o(\mathbf{x}; \omega_o) - \\ \frac{i(\mathbf{x})}{\left[(h(\omega_h) * o(\omega_o)) + \hat{b} \right](\mathbf{x})} \left[h(\mathbf{x}; \omega_h) * \frac{\partial}{\partial \omega_o} o(\mathbf{x}; \omega_o) \right], \end{aligned} \tag{5.20}$$

and

$$\nabla_{\omega_h}\mathcal{J}(\omega_o,\omega_h|i)=\sum_{\mathbf{x}\in\Omega_s}\frac{\partial}{\partial\omega_h}h(\mathbf{x};\omega_h)*o(\mathbf{x};\omega_o)- \\ \frac{i(\mathbf{x})}{\left[(h(\omega_h)*o(\omega_o))+\hat{b}\right](\mathbf{x})}\left[\frac{\partial}{\partial\omega_h}h(\mathbf{x};\omega_h)*o(\mathbf{x};\omega_o)\right]. \quad (5.21)$$

In Eqs. (5.20) and (5.21), we need to calculate the two derivatives $\partial o(\mathbf{x};\omega_o)/\partial\omega_o$ and $\partial h(\mathbf{x};\omega_h)/\partial\omega_h$. These two calculations are summarized below.

Gradient Calculations for the Object Function If the microsphere is not axially centered in the observed volume, then the relative true position $z_0 \neq 0$ and

$$o(\mathbf{x};\omega_o)\sim o(x,y,z-z_o). \quad (5.22)$$

It can be shown that

$$o\left(x,y,z-z_o\right)=o(x,y,z)*\delta\left(0,0,z_o\right), \quad (5.23)$$

$$=\mathcal{F}^{-1}\left\{O(\mathbf{k})\exp\left(-j\frac{2\pi}{N_z}k_z z_o\right)\right\}. \quad (5.24)$$

So, the derivative with respect to the position z_o, $\partial o(\mathbf{x};\omega_o)/\partial\omega_o$, becomes

$$\frac{\partial}{\partial\omega_o}o(\mathbf{x};\omega_o)=\frac{\partial}{\partial z_o}o(x,y,z-z_o), \quad (5.25)$$

$$=-\frac{j2\pi}{N_z}\mathcal{F}^{-1}\left\{k_z O(\mathbf{k})\exp\left(-j\frac{2\pi}{N_z}k_z z_o\right)\right\}. \quad (5.26)$$

By using Eqs. (5.20) and (5.26), we can calculate the gradients with respect to ω_o required at every iteration n of Eq. (5.16).

Gradient Calculation for the Point-Spread Function The calculations for the PSF parameters in Eq. (5.17) are not so straightforward. The gradient of the theoretically calculated PSF in Eq. (3.52) can be written as

$$
\begin{aligned}
\frac{\partial}{\partial \omega_h} h_{Th}(\mathbf{x};\lambda_{\mathrm{ex}},\lambda_{\mathrm{em}}) &= \frac{\partial}{\partial \omega_h}\left\{ |h_A(\mathbf{x};\lambda_{\mathrm{ex}})|^2 \int\limits_{x_1^2+y_1^2\le\frac{D^2}{4}} |h_A(\mathbf{x};\lambda_{\mathrm{em}})|^2 dx_1 dy_1 \right\}, \\
&= |h_A(\mathbf{x};\lambda_{\mathrm{ex}})|^2 \frac{\partial}{\partial \omega_h}\left\{ \int\limits_{x_1^2+y_1^2\le\frac{D^2}{4}} |h_A(\mathbf{x};\lambda_{\mathrm{em}})|^2 dx_1 dy_1 \right\} + \\
&\quad \frac{\partial}{\partial \omega_h}\left\{ |h_A(\mathbf{x};\lambda_{\mathrm{ex}})|^2 \right\} \int\limits_{x_1^2+y_1^2\le\frac{D^2}{4}} |h_A(\mathbf{x};\lambda_{\mathrm{em}})|^2 dx_1 dy_1, \\
&= |h_A(\mathbf{x};\lambda_{\mathrm{ex}})|^2 \int\limits_{x_1^2+y_1^2\le\frac{D^2}{4}} \frac{\partial}{\partial \omega_h}\left\{ |h_A(\mathbf{x};\lambda_{\mathrm{em}})|^2 \right\} dx_1 dy_1 + \\
&\quad \frac{\partial}{\partial \omega_h}\left\{ |h_A(\mathbf{x};\lambda_{\mathrm{ex}})|^2 \right\} \int\limits_{x_1^2+y_1^2\le\frac{D^2}{4}} |h_A(\mathbf{x};\lambda_{\mathrm{em}})|^2 dx_1 dy_1.
\end{aligned}
\tag{5.27}
$$

For Eq. (5.27), it is sufficient to calculate $\partial h_A(\mathbf{x};\lambda)/\partial \omega_h$ for a general wavelength λ and then use the same functions for λ_{ex} and λ_{em}. Thus,

$$
\begin{aligned}
\frac{\partial}{\partial \omega_h}|h_A(\mathbf{x};\lambda)|^2 &= \frac{\partial}{\partial \omega_h}\{h_A(\mathbf{x};\lambda)h_A^*(\mathbf{x};\lambda)\} \\
&= \frac{\partial}{\partial \omega_h}\{h_A(\mathbf{x};\lambda)\}h_A^*(\mathbf{x};\lambda) + h_A(\mathbf{x};\lambda)\frac{\partial}{\partial \omega_h}\{h_A^*(\mathbf{x};\lambda)\} \\
&= \left(h_A^*(\mathbf{x};\lambda)\frac{\partial}{\partial \omega_h}\{h_A(\mathbf{x};\lambda)\}\right) + \left(h_A^*(\mathbf{x};\lambda)\frac{\partial}{\partial \omega_h}\{h_A(\mathbf{x};\lambda)\}\right)^*
\end{aligned}
\tag{5.28}
$$

Assuming the apodization function to be unity, the derivative of the amplitude PSF can be written as

$$\begin{aligned}
\frac{\partial}{\partial \omega_h} h_A(\mathbf{x};\lambda) &= \frac{\partial}{\partial \omega_h} \mathscr{F}^{-1}\{\exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\} \\
&= \mathscr{F}^{-1}\left\{\frac{\partial}{\partial \omega_h} \exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\right\} \\
&= \mathscr{F}^{-1}\left\{jk_0 \exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\frac{\partial}{\partial \omega_h}\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\right\} \\
&= jk_0\mathscr{F}^{-1}\left\{\exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\frac{\partial}{\partial \omega_h}\varphi(\theta_i,\theta_s,z;d,n_i,n_s))\right\}
\end{aligned} \tag{5.29}$$

But,

$$\varphi(\theta_i,\theta_s,z;d,n_i,n_s) = \varphi_d(\theta_s,z;d,n_i,n_s) + \varphi_a(\theta_i,\theta_s,z;d,n_i,n_s), \tag{5.30}$$

$$= n_s z(1-\cos\theta_s) + d(n_s\cos\theta_s - n_i\cos\theta_i). \tag{5.31}$$

It is straight forward to show that

$$\left[\frac{\partial}{\partial \omega_h}\varphi(\theta_i,\theta_s,z;d,n_i,n_s)\right]_{\omega_h=d} = (n_s\cos\theta_s - n_i\cos\theta_i), \tag{5.32}$$

and also

$$\begin{aligned}
\left[\frac{\partial}{\partial \omega_h} n_s\cos\theta_s\right]_{\omega_h=n_s} &= \frac{\partial}{\partial \omega_h} n_s \frac{\left(k_s^2-(k_x^2+k_y^2)\right)^{\frac{1}{2}}}{k_s}, \\
&= \frac{\partial}{\partial \omega_h}\left(n_s^2-n_k^2\right)^{\frac{1}{2}}, \text{ where } n_k = \frac{\lambda}{2\pi}\left(k_x^2+k_y^2\right)^{\frac{1}{2}}, \\
&= \frac{n_s}{\left(n_s^2-n_k^2\right)^{\frac{1}{2}}}, \\
&= \frac{k_s}{\left(k_s^2-(k_x^2+k_y^2)\right)^{\frac{1}{2}}}, \\
&= \sec\theta_s,
\end{aligned} \tag{5.33}$$

where $\sec\theta_s = \cos^{-1}\theta_s$. From the Eq. (5.33), we can say

$$\left[\frac{\partial}{\partial \boldsymbol{\omega}_h}\varphi(\theta_i,\theta_s,z;d,n_i,n_s)\right]_{\omega_h=n_i} = -d\sec\theta_i, \tag{5.34}$$

and

$$\left[\frac{\partial}{\partial \boldsymbol{\omega}_h}\varphi(\theta_i,\theta_s,z;d,n_i,n_s)\right]_{\omega_h=n_s} = z(1-\sec\theta_s)+d\sec\theta_s. \tag{5.35}$$

Summarizing, the gradient of a theoretical PSF model $h_{\mathrm{Th}}(\mathbf{x};\boldsymbol{\omega}_h)$ with respect to its parameters $\boldsymbol{\omega}_h$, can be obtained by combining the Eqs. (5.27), (5.28), (5.29), (5.32), (5.34) and (5.35).

Hessian Calculation Since the Hessian in Eq. (5.16) and (5.17) are difficult to be calculated for each of the parameters, we make an approximation based on the gradients. This is a quasi-Newton method, where an approximation for the Hessian is used instead based on the successive gradient vectors. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) method is one such approach to solve an unconstrained nonlinear optimization problem (*cf.* Nocedal & Wright [1999]).

Object Intensity and Apodization Estimation

We discuss the estimation of the term s, that is a combination of the intensity and the apodization function. We assumed that the apodization function is a constant. Due to strict convexity of the energy function with respect to s, minimizing it in Eq. (5.13) is equivalent to equating the gradient $\nabla_s \mathcal{J}(\hat{\omega}_o,\hat{\omega}_h,s|i)$ to zero. That is,

$$\begin{aligned}\nabla_s \mathcal{J}(\hat{\omega}_o,\hat{\omega}_h,s|i) = &\frac{\partial}{\partial s}\left\{\sum_{\mathbf{x}\in\Omega_s}\left[s(h(\hat{\omega}_h)*o(\hat{\omega}_o))+\hat{b}\right](\mathbf{x}) - \right.\\ &\left.\sum_{\mathbf{x}\in\Omega_s} i(\mathbf{x})\log\left[s(h(\hat{\omega}_h)*o(\hat{\omega}_o))+\hat{b}\right](\mathbf{x})\right\} \qquad (5.36)\\ = &0, \qquad (5.37)\end{aligned}$$

where the terms independent of s was dropped from Eq. (5.13), and the parameters $\hat{\omega}_o$ and $\hat{\omega}_o$ are assumed to be known or *a priori* estimated as describe in the previous subsections.

$$\nabla_s \mathcal{J}(\hat{\omega}_o, \hat{\omega}_h, s|i) = \sum_{\mathbf{x}\in\Omega_s} \left[h(\hat{\omega}_h) * o(\hat{\omega}_o)\right](\mathbf{x}) - \\ \sum_{\mathbf{x}\in\Omega_s} \frac{i(\mathbf{x})}{\left[s(h(\hat{\omega}_h) * o(\hat{\omega}_o)) + \hat{b}\right](\mathbf{x})} \cdot \left[h(\hat{\omega}_h) * o(\hat{\omega}_o)\right](\mathbf{x}) = 0 \quad (5.38)$$

and hence

$$\sum_{\mathbf{x}\in\Omega_s} [h(\hat{\omega}_h) * o(\hat{\omega}_o)](\mathbf{x}) = \sum_{\mathbf{x}\in\Omega_s} \frac{i(\mathbf{x})}{\left[s(h(\hat{\omega}_h) * o(\hat{\omega}_o)) + \hat{b}\right](\mathbf{x})} \cdot \\ \left[h(\hat{\omega}_h) * o(\hat{\omega}_o)\right](\mathbf{x}), \forall \mathbf{x} \in \Omega_s. \quad (5.39)$$

The above equality is possible iff each of the individual terms within the sum are equal.

$$1(\mathbf{x}) = \frac{i(\mathbf{x})}{[s(h(\hat{\omega}_h) * o(\hat{\omega}_o)) + \hat{b}](\mathbf{x})}, \forall \mathbf{x} \in \Omega_s, \quad (5.40)$$

or equivalently,

$$s = \frac{1}{(N_x N_y N_z)} \sum_{\mathbf{x}\in\Omega_s} \frac{i(\mathbf{x}) - \hat{b}(\mathbf{x})}{[h(\hat{\omega}_h) * o(\hat{\omega}_o)](\mathbf{x})}, \forall \mathbf{x} \in \Omega_s\,. \quad (5.41)$$

In an iterative form, the above expression for calculating s can now be written as

$$\hat{s}^{n+1} = \frac{1}{(N_x N_y N_z)} \sum_{\mathbf{x}\in\Omega_s} \frac{i(\mathbf{x}) - \hat{b}(\mathbf{x})}{[h(\hat{\omega}_h^n) * o(\hat{\omega}_o^n)](\mathbf{x})}. \quad (5.42)$$

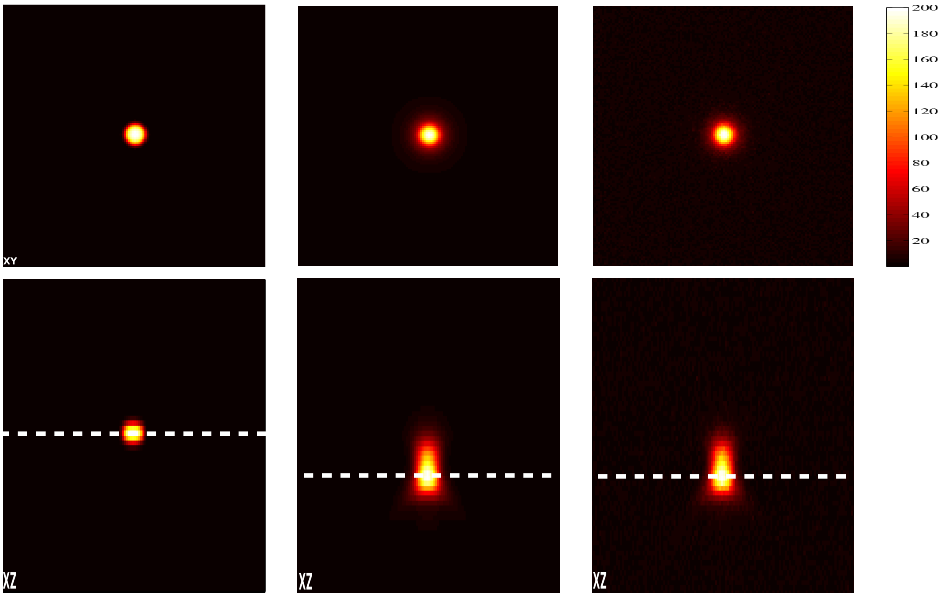
5.3 Implementation and Analysis

5.3.1 Simulating the Observation

For the implementation, a bandlimited object was simulated as explained in Appendix A. The radius of the microsphere for simulations was chosen as 250nm, and is assumed to be embedded in a medium of RI $n_s = 1.33$. Fig. 5.2(a) shows this original object that we wish to image. This object is assumed to be imaged by a Zeiss LSM 510 microscope fitted with a 40X, 1.3 NA '*Plan-Neofluar*' oil immersion lens ($n_i = 1.518$). If the coverslip chosen has a RI very close to that of the objective medium, the only source of aberrations are due to the reason that $n_i \neq n_s$. The excitation and emission peaks are at wavelengths of 488nm (λ_{ex}) and 520nm (λ_{em}) respectively. The LSM pinhole is adjusted to a physical size of 61μm, and the images are Nyquist sampled at lateral and axial pixel sizes of 46.92nm and 166.16nm. For the above imaging setup, the PSF is calculated from Eq. (3.52) at a depth of 5μm, and the object is blurred as in Fig. 5.2(b) with a constant background intensity of 10. The observation in Fig. 5.2(c) is generated with $\gamma = 100$. We denote by $\boldsymbol{\omega}_{h,\mathrm{true}} = \{1.518, 1.33, 10\mu m\}$, the true experimental settings of the simulation. We can clearly notice the shift in axial center of gravity (COG) from z_o in the original object to z_i for its blurred version. This axial shift was discussed in Subsection 3.4.1. As mentioned before, we recall once again that the radial center for the object (x_o, y_o) and that in the image plane (x_i, y_i) remains unchanged.

5.3.2 Initialization of the Algorithm

As the microsphere is not centered, its approximate initial relative position $(x_o^{(0)}, y_o^{(0)}, z_o^{(0)})$ has to be calculated from the observed images. We propose a simple approach in estimating the relative microsphere position in the observed images. The only assumption made here is that the observation data has been treated so that the volume is from a single microsphere. By knowing the physical diameter of the microsphere (here 500nm), we can locate its initial position in the observed volume by estimating the COG of the intensities (*cf.* Peng, et al. [2007]). We remind that the lateral object position does not change during the observation process, and hence the estimation of the lateral centroid of the observed microsphere

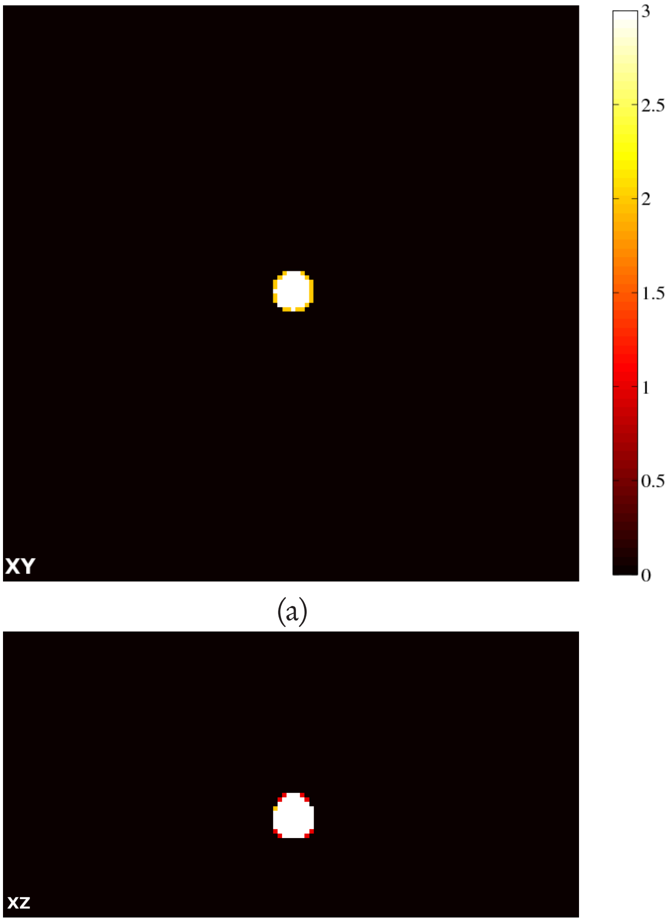


(a) (b) (c)

Figure 5.2: Simulation of a (a) microsphere object, (b) the blurred observation showing the background fluorescence $b = 10$, (c) the blurred observation with the background fluorescence and Poisson noise with $\gamma = 100$. The object has a maximum intensity of 200IU and is Nyquist sampled at 46.92nm and 166.16nm. (©Ariana-INRIA/I3S).

corresponds to the true radial microsphere position. The above method gives very precise lateral locations $(\hat{x}_o^{(0)}, \hat{y}_o^{(0)})$. For example, Fig. 5.3 shows the segmentation of the observation data in Fig. 5.2(c) by using the above method. However, due to focal anomaly the axial relative location $(\hat{z}_o^{(0)})$ of the object cannot be accurately obtained from the observation data. In the example given above, the COG of the observation, the nominal position was accurately estimated as 8 slices off the central plane. However, this does not correspond with the actual position of the object in the volume.

From geometrical optics (*cf.* Visser, et al. [1992]), if the objective is non-paraxial, the estimated axial position from the observation, $\hat{z}_o^{(0)}$, is multiplied by a factor $(\tan(\arcsin(NA/n_i^{(0)}))/(\tan(\arcsin(NA/n_s^{(0)})))$ to get the new relative position $(\hat{z}_{o,\text{new}}^{(0)})$. While for the paraxial case, as the angles are small, the tangent of the angles can be approximately equal to the sine of the angle, and thus the multiplication factor becomes $(n_s^{(0)}/n_i^{(0)})$. These new values are then used as starting points for the iteration. It was found that the initialization process had a



(a)

(b)

Figure 5.3: Segmentation of the MIP along the (a) optic axis projection giving radial segment and (b) lateral projection giving axial segment (©Ariana-INRIA/I3S).

maximum error of 1% off the true lateral position, and about 8.4% off the true axial position. The error in the axial position is significant and hence it is not correct to use these values directly as the true position. The parameters of the PSF $\hat{\omega}_h^{(0)}$ are assigned valid non-zero values as starting estimates. The refractive indices are sampled from the interval $[0, 1.5]$ so that $n_i^{(0)} \neq n_s^{(0)}$. The initial depth is chosen as any value greater than zero.

5.3.3 Preliminary Results on Simulated Data

In Fig. 5.4, we show the ℓ_∞ normalized axial intensity profiles for the true object and the observation. The intensity profiles for two deconvolution results is also plotted. In the first profile, the observation is deconvolved with a defocus PSF under no aberrations. While the second profile is the deconvolution with the correct SA PSF. It can be seen that the observed intensity profile is axially shifted

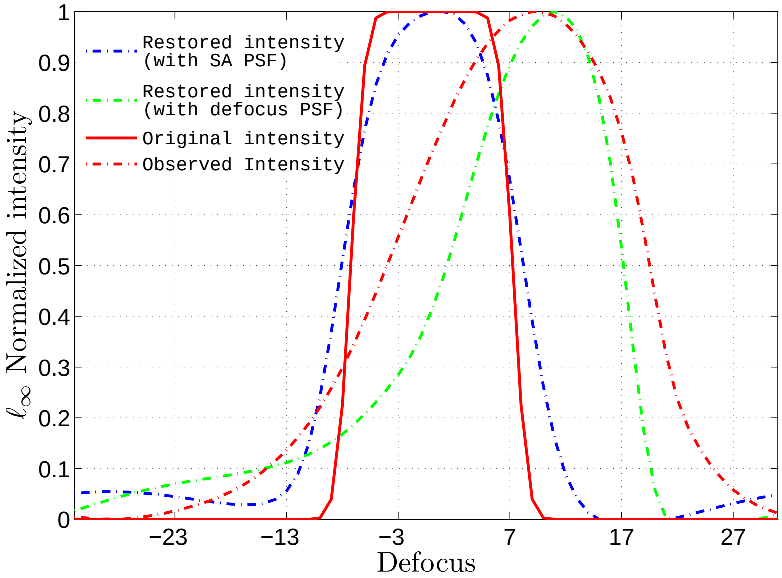
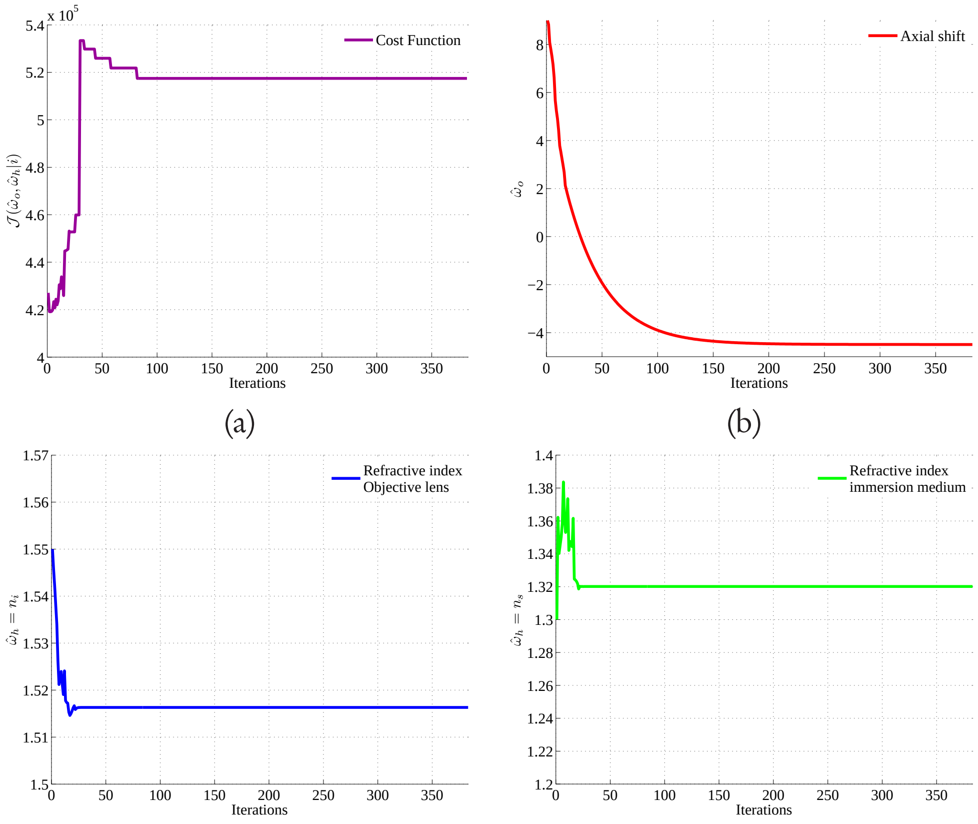


Figure 5.4: Axial intensity profiles for the true object, the observed object, the restored intensity with a diffraction-limited PSF, and the restored intensity with a spherically aberrated PSF. The x-axis gives the z-plane numbers and the intensities are ℓ_∞ normalized for visual comparison (©Ariana-INRIA/I3S).

from the true axial position by about 8 slices. After restoration with the SA PSF, the estimated object's axial location corresponds very well with the true position of the object. The intensity profile is quite symmetrical in comparison to the observed profile. However, when restoration is done using a diffraction-limited approximation (ignoring aberrations), it led to error in the actual axial position. We can also observe that the intensity is highly asymmetrical and is worse than the observation profile.

We also discuss the results obtained by using the JMAP algorithm proposed in Subsection 5.2.2. We follow the iteration progress for the cost function Eq. 5.11 and the different parameters for the simulation in Fig. 5.5. The refractive indices n_i and n_s were estimated to be 1.5163 and 1.3201 which are accurately close to their true values of 1.518 and 1.33 chosen for the simulation. Although these two parameters were accurately determined with a maximum error of $< 1\%$, there were errors in estimating the relative true position of the object in the volume



(a) (b)

(c) (d)

Figure 5.5: Graph showing the progression of the (a) cost function $\mathcal{J}(\hat{\omega}_o,\hat{\omega}_h|i)$, (b) relative position $\hat{\omega}_o$, (c) objective lens refractive index $\hat{n}_i$, and (d) specimen medium refractive index $\hat{n}_s$ estimation with iterations (ⓒAriana-INRIA/I3S).

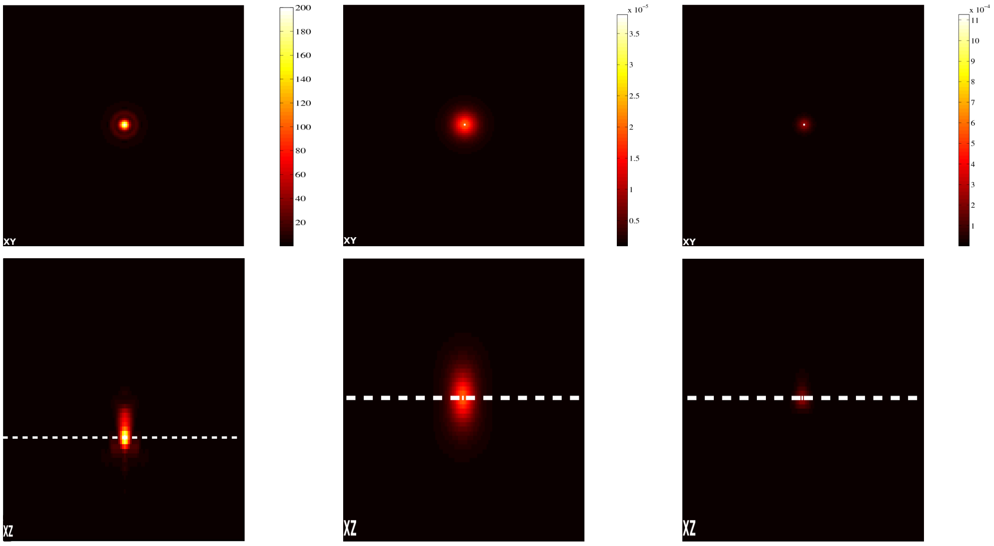
and the NFP d. The relative position of the object in the volume was found to be about 4 slices below the central plane, and differs from the true position by about 2 slices. Due to this reason, the cost function does not reach its global minimum as seen in Fig. 5.5(a). This problem could be a result of the approximation we used during optimization and hence could be overcome by choosing the initial depth $d^{(0)}$ closer to its true value.

We compare the results obtained with the naïve MLEM BD algorithm. In Fig. 5.6, we show the results of applying Eqs. (2.34) and (2.35) to the observation of Fig. 5.2(c). As this algorithm has to be be manually terminated, we show the results for two iterations 70 and 200, with $q_o = 1$ and $q_h = 1$. Fig. 5.6(a) and (d) shows the true SA PSF and the true object used for simulating the observation. Since the algorithm does not have any information about the object or the PSF, the estimate of the object $\hat{o}^{(n)}(\mathbf{x})$ in Fig. 5.6(e) and (f), resembles the true PSF, Fig. 5.6(a), in shape and in the position of the COG. Similarly, the estimate of the PSF, $\hat{h}^{(n)}(\mathbf{x})$, in Fig. 5.6(b) and (c) resembles more the imaged microsphere, Fig. 5.6(d), that is axially centered. Thus, having no prior knowledge on the object or the PSF can make it difficult to distinguish them. The characteristics of the PSF is absorbed by the estimated object and vice versa. Also, when we do not stop the iterations, progressively the results start deteriorating with the algorithm favoring a Dirac beyond 200 iterations of the naïve MLEM BD. We perform another experiment, where the object is axially shifted from its centroid by two planes, as in Fig. 5.7(a). Fig. 5.7(c) shows the same PSF used in Fig. 5.6(a). Fig. 5.7(b) and (d) show $\hat{h}^{n}(\mathbf{x})$ and $\hat{o}^{n}(\mathbf{x})$ with $n = 70$ iterations. This experiment leads us to conclude that the true axial position of the object cannot be recovered using this naïve MLEM BD.

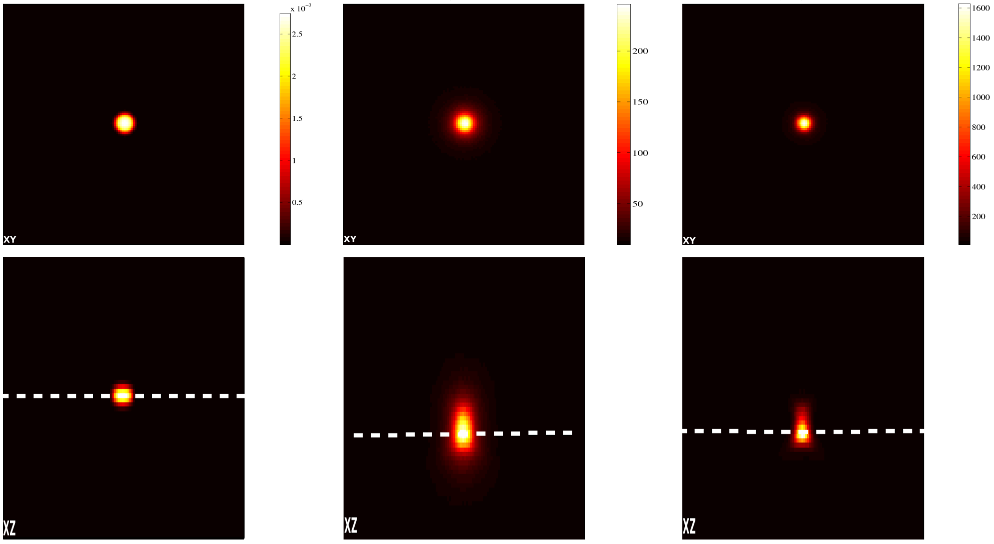
5.3.4 Empirically derived Point-Spread Function

There are several methods available in literature for experimentally imaging microspheres notable ones are by Gibson & Lanni [1989]; Hiraoka, et al. [1990]. However, for our experiments, we are interested in artificially generating a condition where there is an intentional mismatch in refractive indices between the objective lens and the specimen medium. In order to achieve this, we used a very simple imaging setup. Some fluorescent microspheres are stuck to the bottom of a coverslip, and the coverslip is placed in water. In this way the depth is fixed (thickness of the coverslip), and there are no additional aberrations. Fig. 5.8 shows the schematic of the experiment. If the immersion medium is either water or glycerol, SA appears due to mismatch in index between n_i and n_g.

The use of very small spheres lead to bad SNR, so one often uses spheres of the order of the microscope resolution. We used polystyrene latex microspheres from Tetraspeck$^{\mathrm{TM}}$with the manufacturer specified diameter of about 170nm. These spheres have a peak excitation/emission wavelengths of 505/515nm (green). The coverslip is of type 1.5, and has a thickness of 170μm and RI of about 1.522. The objective used is a '*C-Apochromat*' water immersion lens of 63X magnification and NA 1.2. The pinhole is set at a physical size of $112\mu m$. The imaged microspheres are larger than the recommended size for them to be considered as point sources. However, this does not pose a problem as we have a method for



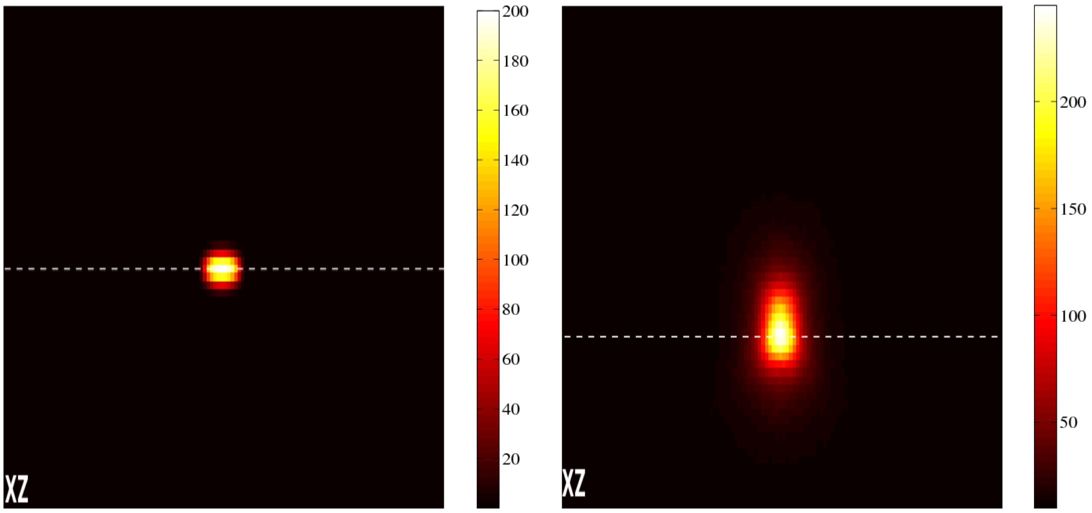
(a) True PSF $h(\mathbf{x})$ (b) $\hat{h}^{(n)}(\mathbf{x})$, $n = 70$ (c) $\hat{h}^{(n)}(\mathbf{x})$, $n = 200$



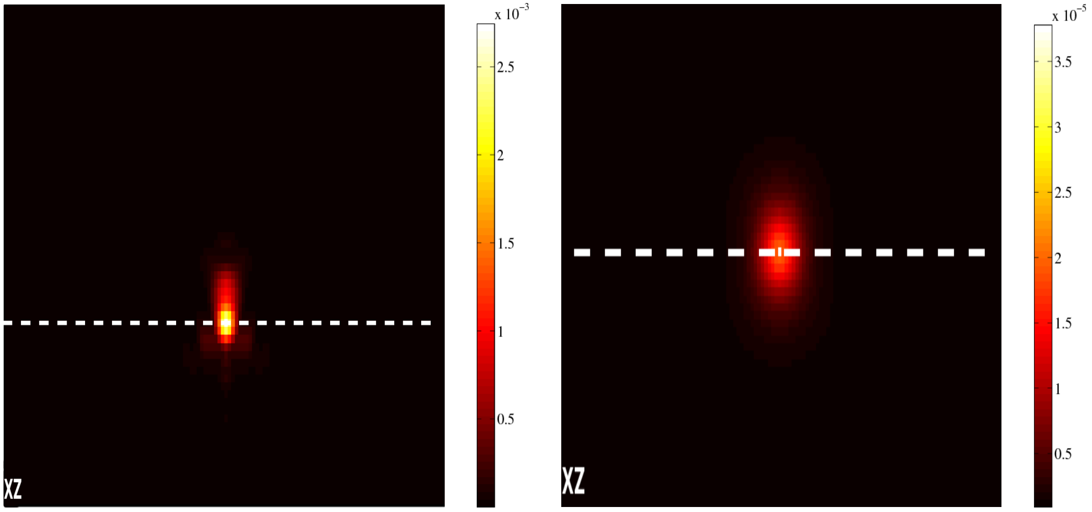
(d) true object $o(\mathbf{x})$ (e) $\hat{o}^{(n)}(\mathbf{x})$, $n = 70$ (f) $\hat{o}^{(n)}(\mathbf{x})$, $n = 200$

Figure 5.6: Blind deconvolution results with the naïve MLEM algorithm after 70 and 200 iterations (©Ariana-INRIA/I3S).

compensating this in the estimation procedure. Fig. 5.9 shows the microspheres distributed at various positions in the radial plane. Since they are all stuck to the bottom of the cover slip, they are at the same depth. The images are sampled at a radial sampling for 0.037μm and axial sampling of 0.151μm. From the ZY section, we see that two spheres that are positioned at the same depth but at different radial position are very similar. This validates our assumption that the PSF is approximately invariant to radial translations. Fig. 5.10(a) shows the radial and



(a) True object $o(\mathbf{x})$, (b) estimated object $\hat{o}^{(n)}(\mathbf{x})$



(c) true PSF $h(\mathbf{x})$ (d) estimated PSF $\hat{h}^{(n)}(\mathbf{x})$

Figure 5.7: (a) Simulated true object $o(\mathbf{x})$ shifted along the optic axis of a volume, (b) the estimated object $\hat{o}(\mathbf{x})$ using the naïve MLEM BD, (c) the true PSF $h(\mathbf{x})$ used for simulation, and (d) the estimated PSF $\hat{h}(\mathbf{x})$ using the naïve MLEM BD. The iterations were manually terminated after $n = 70$ (ⓒAriana-INRIA/I3S).

axial maximum intensity projection (MIP) of the cropped observation of a single 170nm microsphere stuck on to a 1.5 type cover slip. Since the objective used was a water immersion lens, and we are imaging into a cover slip, the resulting observed PSF was axially asymmetrical. These aberrated bead volumes were then radially averaged, and Fig. 5.10(b) shows the corresponding MIP of the radial and axial planes. This resulting radially averaged images could be used to test the al-

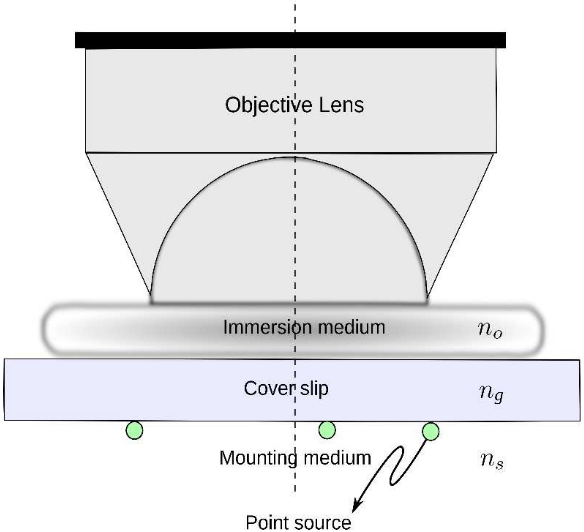


Figure 5.8: The experimental set-up schematic for generating spherical aberration due to refractive index mismatch when imaging point sources (©Ariana-INRIA/I3S).

gorithms on PSF retrieval.

In Fig. 5.10(c), the radially average images were subjected to the VST in Eq. 2.19, denoised using a TV scheme (Rudin, et al. [1992]) and then transformed back. The denoising was individually performed on the 2D sections and they were then stacked together. From these images, it can be seen that the effect of denoising is to smooth the finer details of the PSF. Hence, this procedure is not recommended to retrieve the PSF from the empirical images.

5.4 Conclusion

In this chapter, we propose an approach for estimating the PSF from an observation data given some knowledge of the object. The validation of the algorithm on simulated data shows very promising results for the problem of PSF extraction from observed intensities for a fluorescence microscope where SA is the dominant form of aberration. In the case where sub-resolution microsphere is used, sphere size correction will not be necessary and the object should be treated as a Dirac. Future work is aimed at testing the proposed approach on images of fluorescent polystyrene latex Tetraspeck$^{\text{TM}}$microspheres from Invitrogen$^{\text{TM}}$. One possible extension of this work might involve applying this method to restoring

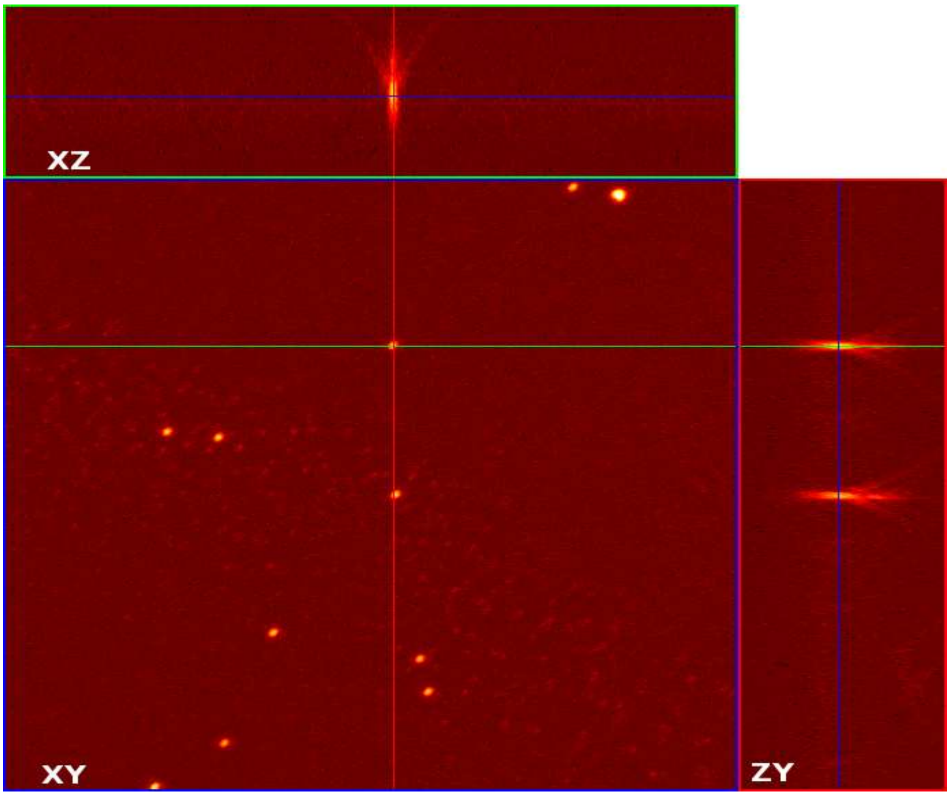


Figure 5.9: Observed microspheres that are stuck to the bottom of a coverslide and distributed along the radial plane (ⓒINRA Sophia-Antipolis). The objective used is a C-Apochromat water immersion lens with a magnification of 63X and NA 1.2. $\Delta_{xy} = 0.037\mu$m and $\Delta_z = 0.151\mu$m.

images of biological specimens affected by spherical aberrations adding some constraints on the object (spatial or frequency). The task is not simple as there are many possible solutions for the phase function, though a realization might be possible through regularization. This work also opens up new possibilities into the field of depth-varying image restoration.

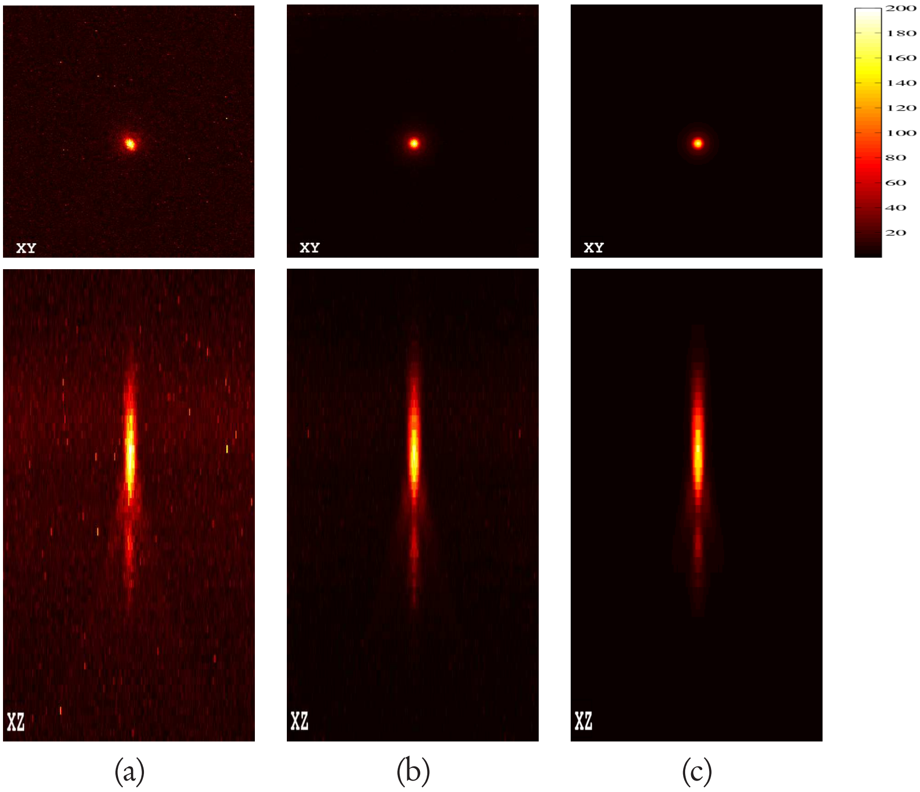


Figure 5.10: Maximum intensity projection along the (a) optic axis gives the lateral plane (top) and radial axis gives axial plane (bottom) of an observed microsphere; (b) the planes after circular averaging the data (a); (c) after denoising the circularly averaged data in (b). The z-sections are purposefully enlarged for ease of viewing (©Ariana-INRIA/I3S, INRA Sophia-Antipolis).

CHAPTER 6

Perspectives on Blind Deconvolution for Thick Specimens

> “*You cannot depend on your eyes when your imagination is out of focus.*”
>
> *-Mark Twain (American humorist, writer and lecturer)*

In this chapter, we look at the prospects of extending the phase and object estimation procedure derived in Chapter 5 for restoring observed specimens. When restoring thick sections ($> 1\mu$m) of specimens, the effects of spherical aberration becomes prominent. This is because the differences in refractive indices is significant with depth below the coverslide. The linear space invariant approximation for the convolution operation in the model of Eq. (1.8) is no longer valid. When handling space varying PSFs, we have to provide a new quasi-convolution observation model. We remodel the blurring equation of Eq. 1.8 by using this quasi-convolution operation and proposed methods to estimate the object from spherically aberrated observation.

6.1 Introduction

Oil immersion lenses are designed for imaging specimens tightly pressed to the coverslip. If a medium appears between the coverslip and the specimen or if imaging through more than 20μm of cell volume, the light-focusing properties of the oil immersion lens begins to falter (*cf.* Shaw [2006]), unless the specimen has the same RI as that of oil. For example, when imaging biological tissues, there is a dramatic reduction in both the signal level and the resolution. Booth, et al.

[1998] mentions that when focusing into water, using an oil immersion objective, the signal level falls to 40% when focusing only $5\mu m$ beneath the coverslip and is below 10% at $15\mu m$. Török, et al. [1997] reviewed a number of theories for describing specimen-induced SA in confocal microscopy. Török obtained numerical results specifically for focusing though a stratified specimen medium. The role of the specimen induced aberration is so significant, that it was suggested to use as thin a sample as possible to minimize this effect. Practically, this poses a lot of difficulty to biologists as more time needs to be invested and also this removes all the advantages gained by optical sectioning. Thus, restoration of thick samples becomes important and hence this chapter is targeted at BD using depth-dependent PSFs.

In Fig. 6.1, we show an optical setup when imaging with oil immersion lenses and with the specimen in water (*cf.* Kam, et al. [2007] for setup with AO correction). As the water/specimen index of refraction is different from the index of the objective immersion medium (say oil or glycerol), there is a rapid decrease in resolution, contrast and peak intensity of the image. These degradations are due to depth dependent SA. It can be theoretically shown that for a given object function $o(\mathbf{x})$ and for fixed acquisition settings (n_i, n_s), the variation of the PSF

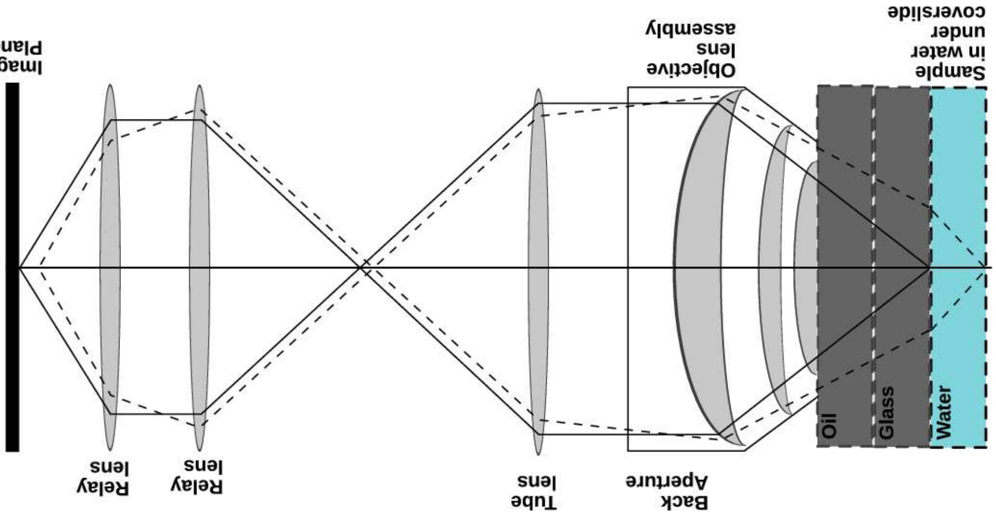


Figure 6.1: Schematic of the optical set-up for spherical aberrations when an oil immersion objective is used for imaging water embedded sample. The solid lines shows rays focus under the coverslide requiring no aberration correction while dashed lines shows focus change as a result of sample in water (©Ariana-INRIA/I3S).

$h(\mathbf{x};\omega_h)$ and the blurred object $\big(h(\mathbf{x};\omega_h)*o(\mathbf{x})\big)$ with the depth d is such that,

$$||h(\mathbf{x};d=0)||_2^2 > ||h(\mathbf{x};d\neq 0)||_2^2\,,\ \text{and} \tag{6.1}$$

$$||(h(\mathbf{x};d=0)*o(\mathbf{x}))||_2^2 > ||(h(\mathbf{x};d\neq 0)*o(\mathbf{x}))||_2^2\,. \tag{6.2}$$

As explained in Chapter 3, their ℓ_1 norms satisfy the following constraints:

$$||\,h(\mathbf{x};d=0)||_1 = ||h(\mathbf{x};d\neq 0)\,||_1\,,\ \text{and} \tag{6.3}$$

$$||\,(h(\mathbf{x};d=0)*o(\mathbf{x}))\,||_1 = ||\,(h(\mathbf{x};d\neq 0)*o(\mathbf{x}))\,||_1\ . \tag{6.4}$$

In Fig. 6.2, we computed the ratio of the ℓ_2 norms between $||(o(\mathbf{x})*h(\mathbf{x};d))||_2^2$ and $||(o(\mathbf{x})*h(\mathbf{x};d=0))||_2^2$ for different depths, and for a lens with $\mathrm{NA}=1.4$. In

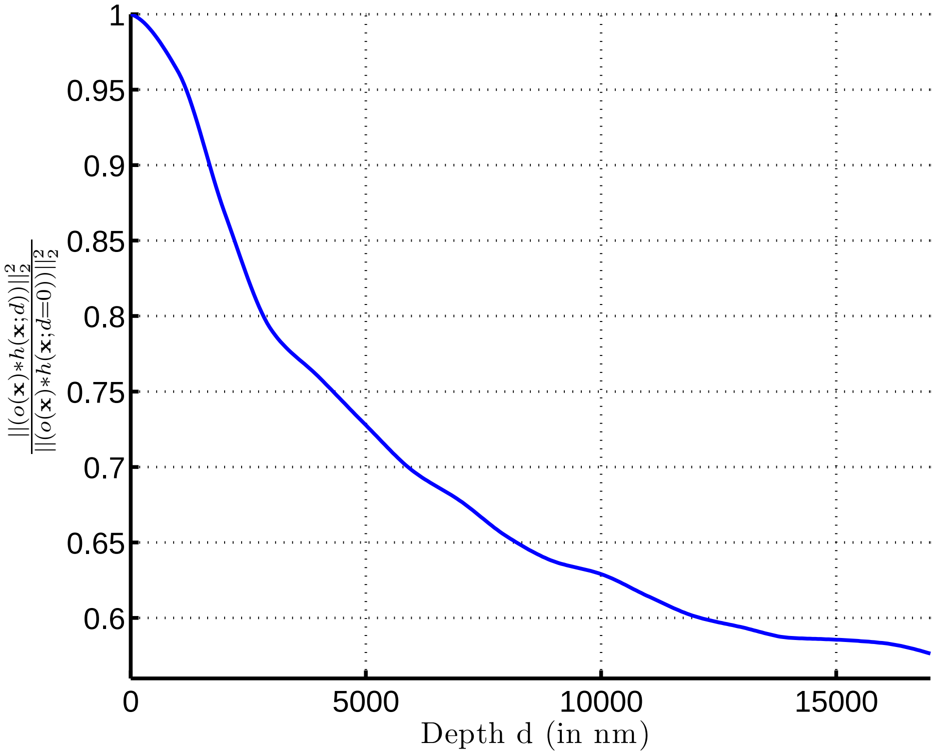


Figure 6.2: For a given simulated object $o(\mathbf{x})$, the plot shows the ratio between the ℓ_2 norm of an observation affected by spherical aberrations and the observation under diffraction-limited case. The NA of the objective lens here is assumed to be 1.4. For low NA lenses the change with depth is minimum (©Ariana-INRIA/I3S).

the literature of optics, this ratio is similar to the Strehl's ratio (*cf.* Born & Wolf [1999]), which calculates the proportion of observed peak intensity at the detection plane from a point source to the theoretical maximum peak intensity of a perfect imaging system working at the diffraction limit. If a lower NA objective (say 0.2) is used, the amount of aberration produced is also lower and it is possible to image deeper sections of the specimen. If the outer part of the back focal plane (BFP) of an oil immersion lens is not filled with plane laser light, it behaves as though the aperture is lower, and the SA produced will also be lower. In Kam, et al. [2007], it was mentioned how when focusing into water, with a high NA lens, this ratio deteriorates rapidly with depth. The AO element that corrects for the depth aberration, was able to greatly minimize the degradation of the Strehl's ratio. We shall see in this chapter how this aberration correction could also be accomplished by employing methods that involve processing the observed images.

Conventional non-blind and blind deconvolution algorithms cannot be used in the presence of aberrations. This is because the assumption of depth invariance is no longer justified. The aberrations in the observed sample renders the conventional deconvolution methods useless. In this chapter, we will focus on image restoration for biological specimens in the presence of these aberrations. We propose a new approach that handles the problem of space varying PSF by using a quasi-deconvolution algorithm.

6.2 Image Formation under Aberrations

6.2.1 Quasi Convolution Model

As the PSF changes with the depth, the assumption of space invariance in Eq. (1.8) is not justified anymore. If the observed image is $i(\mathbf{x})$, its mean $g(\mathbf{x})$ is written as a quasi-convolution in the continuous domain by the following superposition integral (Preza & Conchello [2004]):

$$g(\mathbf{x}) = \int\limits_{\mathbf{x}' \in \Omega_s} h(\mathbf{r}-\mathbf{r}', z; z', \Delta n_s) o(\mathbf{r}', z') d\mathbf{x}' , \tag{6.5}$$

where $\mathbf{x}' = (x', y', z')$ and $\mathbf{r}' = (x', y')$ are the 3D and 2D coordinates in the object space. Similarly, $\mathbf{x} = (x, y, z)$ and $\mathbf{r} = (x, y)$ are the 3D and 2D coordinates in the image space. As the PSF varies with the depth, the calculation of Eq. (6.5) becomes complicated with a requirement for different 3D PSFs for each plane of the observation. In order to overcome this problem, we can use the modified observation model by Preza & Conchello [2004]; Bardsley, et al. [2006], with the object space subdivided into non-overlapping strata. In this manner, the PSF is computed only for fixed depths and is assumed to be invariant throughout a stratum. A similar approach was used by Maalouf, et al. [2008] for space-varying restoration but using a numerically calculated PSF.

As in Eq. 2.16, the data $i(\mathbf{x})$ is assumed to be drawn from an independent Poisson distribution with mean $g(\mathbf{x})$ so that the likelihood becomes

$$\Pr(i(\mathbf{x})|o(\mathbf{x}),\omega_h) = \prod_{\mathbf{x}\in\Omega_s} \frac{(g(\mathbf{x}))^{i(\mathbf{x})}\exp(-g(\mathbf{x}))}{i(\mathbf{x})!}\,, \tag{6.6}$$

where $\omega_h = \{\Delta n_s\}$. In the presence of any stray external fluorescence, $g(\mathbf{x})$ also includes the background term $b(\mathbf{x})$ (*cf.* Subsection 1.3.2).

6.2.2 Influence of Refractive Index

We reconsider the pupil function in Eq. (3.28) of Chapter 3 with the aberrations as

$$P(k_x,k_y,z) = \begin{cases} \exp(jk_0\varphi(\theta_i,\theta_s,z;d,n_i,n_s)), & \text{if } \frac{\sqrt{k_x^2+k_y^2}}{k_i} < \frac{NA}{n_i} \\ 0, & \text{otherwise.} \end{cases} \tag{6.7}$$

We ignore the apodization function and focus on the description of the pupil using the phase $\varphi(\theta_i,\theta_s,z;d,n_i,n_s)$. It was shown in Eq. (3.49) that the phase due to SA could be approximated as a function of the depth and the indices difference, $\Delta n_s = (n_s - n_i)$. Thus, the phase is

$$\begin{aligned} \varphi(\theta_i,\theta_s,z;d,n_i,n_s) &\approx \varphi_d(\theta_i,z;n_i) + \varphi_a(\theta_s;d,n_i,\Delta n_s) & (6.8)\\ &= n_i z(1-\cos\theta_i) + d\Delta n_s \sec\theta_s\,, & (6.9) \end{aligned}$$

and the amplitude PSF of Eq. 3.27 is simply written as $h_A(\mathbf{x};d,\Delta n_s)$. As a result, for a given lens system, the light distribution in the observed images is only affected by the relative change in the RI, and by the NFP. For example, when using oil immersion lenses ($n_i = 1.518$), the distortions for imaging a point source at a depth of 100μm in glycerol solution ($\Delta n_s = 0.0434$) is similar to that of imaging a point source at a depth of 23μm in water ($\Delta n_s = 0.188$). In Fig. 6.3, we numerically computed the PSFs using the approximations derived in this chapter. Note how the asymmetry of the PSF along the optic axis changes between the oil immersion lens (negative SA) and the water immersion lens (positive SA), and with the specimen medium.

The primary source of sample induced distortions is due to the RI variations within the sample (*cf.* Török, et al. [1997]). However, the PSF cannot be numerically computed if this variation is unknown. In Kam, et al. [2001], a ray tracing method was proposed to obtain the PSFs, while in Simon, et al. [2008], a combination of microholography and tomographic illumination was used to obtain this variation. Once this information is known, these aberrations could be compensated from the observation. We will see how Bayesian inference could also be used to estimate this difference, Δn_s, from the observed images.

6.3 Refractive Index Variation and Object Estimation by Bayesian Inference

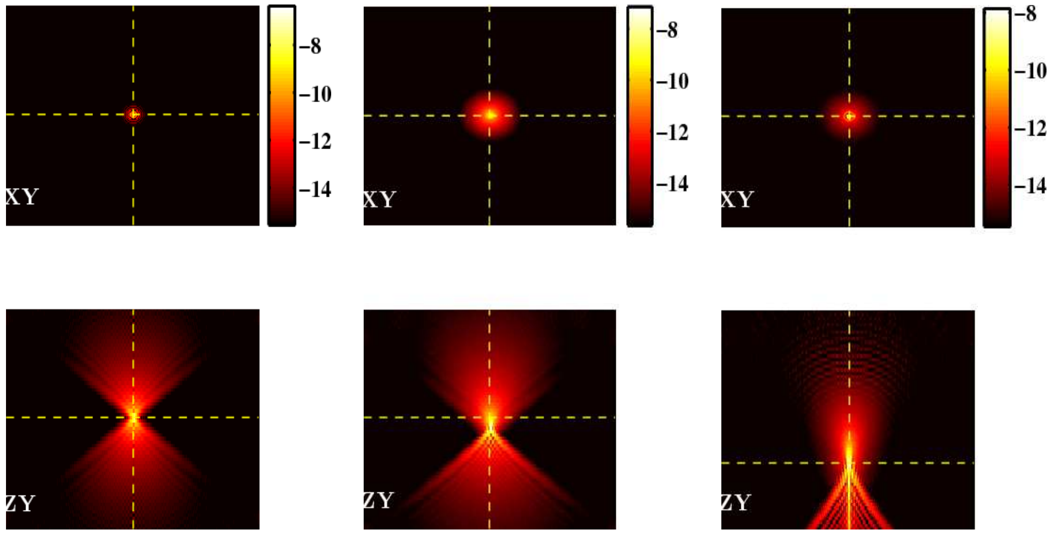
Based on our previous study on phase retrieval by parameter estimation (*cf.* Chapter 5), we propose to recover both the object and the PSF by using the Bayesian inference. Directly maximizing Eq. (6.6) gives the following familiar joint ML estimate:

$$(\hat{o}, \hat{\omega_h}) = \underset{(o \geq 0, \omega_h \in \Theta)}{\operatorname{argmax}} \left\{ \Pr(i | o, \omega_h) \right\}, \tag{6.10}$$

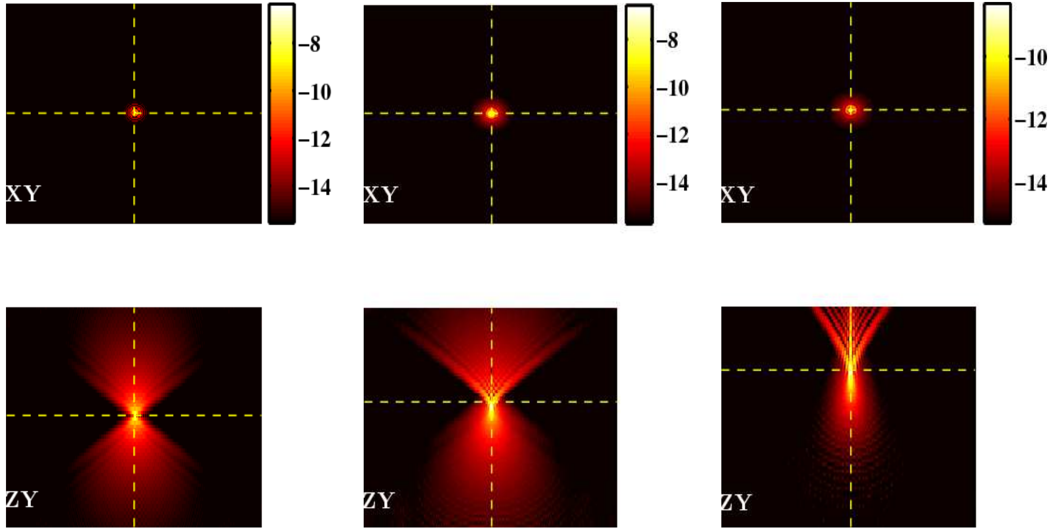
$$= \underset{(o \geq 0, \omega_h \in \Theta)}{\operatorname{argmin}} \mathscr{J}_{\text{obs}}(i(\mathbf{x}) | o(\mathbf{x}), \omega_h) \tag{6.11}$$

We get the cost function $\mathscr{J}_{\text{obs}}(i(\mathbf{x}) | o(\mathbf{x}), \omega_h)$ by applying the $-\log$ operator on the likelihood as

$$\mathscr{J}_{\text{obs}}(i(\mathbf{x}) | o(\mathbf{x}), \omega_h) = \sum_{\mathbf{x} \in \Omega_s} g(\mathbf{x}) - i(\mathbf{x}) \log g(\mathbf{x}), \tag{6.12}$$



(a) $d = 0\mu$m, $n_i = 1.518$ b) $d = 5\mu$m, $n_i = 1.518$ (c) $d = 30\mu$m, $n_i = 1.518$



(d)$d = 0\mu$m, $n_i = 1.33$ (e) $d = 5\mu$m, $n_i = 1.33$ (f) $d = 30\mu$m, $n_i = 1.33$

Figure 6.3: The numerically computed widefield microscope's PSFs using the approximation in Eq. (6.9), for variations in depth and refractive indice. The PSFs in (a)-(c) are simulated for a point source in water ($n_s = 1.33$) and imaged with oil immersion objective lenses ($n_i = 1.518$), while the PSFs in (d)-(f) are for the point source in glycerol ($n_s = 1.4746$) and imaged with water immersion lens ($n_i = 1.33$). The depth under the coverslide is varied as 0μm, 5μm, 30μm (ⓒAriana-INRIA/I3S).

where $g(\mathbf{x})$ is as given in Eq. (6.5).

6.3.1 Joint Maximum *a Posteriori* Estimate

Since the object and index estimation are both ill-posed problems, we add a regularization on the object as in Chapter 4 and the on the index difference, Δn_s. The regularization of the object could be accomplished by using the TV functional as in Eq. (4.6) with a positive object flux. While for the RI, we assume it to be Gaussian distributed so that

$$\Pr(\omega_h) \sim \exp\left(-\frac{1}{2}\left(\omega_h-\mu_{\omega_h}\right)^T \Sigma^{-1}\left(\omega_h-\mu_{\omega_h}\right)\right), \tag{6.13}$$

where μ_{ω_h} is the *a priori* difference mean, and Σ is the covariance matrix. The regularization energy function can be written as

$$\mathcal{J}_{\text{reg},\omega_h} = \frac{1}{2}\left(\omega_h-\mu_{\omega_h}\right)^T \Sigma^{-1}\left(\omega_h-\mu_{\omega_h}\right). \tag{6.14}$$

The combined cost function to be minimized can be written as

$$\begin{aligned}\mathcal{J}(o(\mathbf{x}),\omega_h|i) &= \mathcal{J}_{obs}(i|o(\mathbf{x}),\omega_h)+\mathcal{J}_{\text{reg},o}(o(\mathbf{x}))+\mathcal{J}_{\text{reg},\omega_h}(\omega_h), && (6.15)\\ &= \sum_{\mathbf{x}\in\Omega_s} g(\mathbf{x})-i(\mathbf{x})\log g(\mathbf{x})+\lambda_o\sum_{\mathbf{x}\in\Omega_s}|\nabla o(\mathbf{x})|+ \\ &\quad \frac{1}{2}\left(\omega_h-\mu_{\omega_h}\right)^T \Sigma^{-1}\left(\omega_h-\mu_{\omega_h}\right) && (6.16)\end{aligned}$$

The minimization of Eq. (6.16) could be achieved by using the AM algorithm of Chapter 4, which gives the MAP estimates for the two functions as

$$\hat{o}^{(n+1)} = \underset{(o\geq 0)}{\operatorname{argmin}}\,\mathcal{J}(o(\mathbf{x}),\hat{\omega}_h^{(n)}|i(\mathbf{x})), \text{ followed by} \tag{6.17}$$

$$\hat{\omega}_h^{(n+1)} = \underset{(\omega_h\in\Theta)}{\operatorname{argmin}}\,\mathcal{J}(\hat{o}^{(n)}(\mathbf{x}),\omega_h|i(\mathbf{x})). \tag{6.18}$$

The minimization in Eq. (6.17) could be achieved by a modified MLEM, as in Preza & Conchello [2004], but with TV as the regularization. In Chapter 5, we had discussed a quasi-Newton approach, BFGS, for estimating the parameters of the phase and the object. The estimation of the index variation could be accomplished using the same approach, but with the assumption that, in the sample, it is slowly varying.

6.4 Future Directions

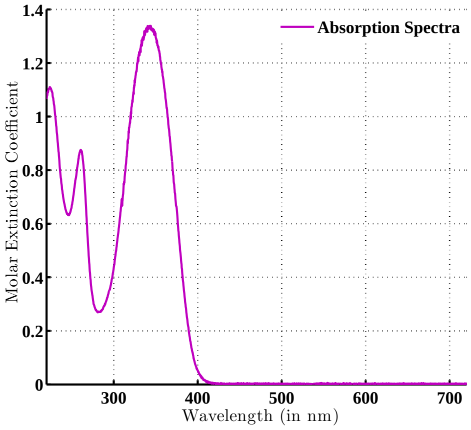
6.4.1 Multichannel Estimation

The aim of adding regularization during the deconvolution is to make the problem well-posed. In fluorescence microscopy, it is very difficult to get two successive scanned observations of the same sample. Regularization is a way of bypassing this need to provide more observations of the same sample. As mentioned earlier, the method of structured illumination overcomes this problem by superimposing two or three different laterally shifted images in the Fourier domain.

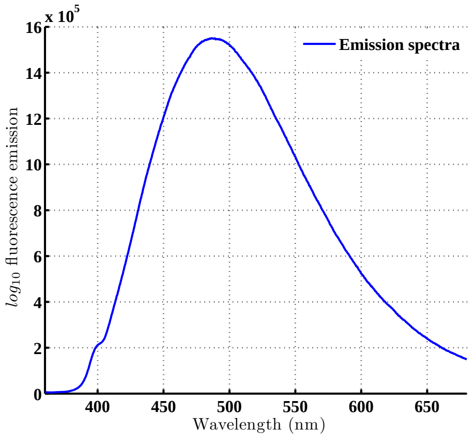
4',5-diamidino-2-phenylindole or DAPI stain is used for tagging specimens, and is usually excited using a ultra violet laser (*cf.* Fig. 6.4(b); data source: Du, et al. [1998]). As the emission for this stain is fairly broad, the signal is observed as blue/cyan in the range 400 to 600nm (*cf.* Fig. 6.4(b); data source: Du, et al. [1998]). The peak appears at a wavelength of 461nm. By using different filters, it is possible to acquire two different sets of images simultaneously based on differences in the ranges of wavelengths (*e.g.* $390-465$nm and $565-615$nm). We have thus at our disposal, two images observed from the convolution of the same object but with two different PSFs. Thus, the information of a given object is increased by using multichannel observations. A similar 2D blind deconvolution method for objects blurred with multiple FIR filters was demonstrated by Harikumar & Bresler [1999].

6.4.2 Discussion

In this chapter, we proposed an approach to restore observations affected by spherical aberration due to mismatch in the refractive index between objective immersion medium and specimen medium. In the presence of SA, the LSI model is not applicable, and a new quasi-convolution model is proposed. This model divides the object space into non-overlapping strata, so that the algorithm in Eq. 4.9 can be run piece-wise over the whole volume (*cf.* Preza & Conchello [2004]). Bayesian inference can be used to obtain the PSF by estimating the RI variation over the volume. This work, thus opens up new avenues in space-varying restoration and the study of specimen RI variation. Future work is aimed at testing this approach on simulated and real images. The approach presented in this chapter is



(a)



(b)

Figure 6.4: (a) Graph of molecular extinction coefficient, and (b) the fluorescence spectrum of DAPI dissolved in water. Data source: see text.

not restricted to CLSM or WFM, as there are no constraints on the image space but only on the pupil phase. Thus, it could be extended to other newer imaging techniques as well.

APPENDIX A

Simulating a Band-Limited Object

Since we frequently use simulated data set for testing or comparing the performance of our algorithm, it is necessary to numerically approximate a band-limited object that could be used to represent a real object that we are interested in imaging. The geometry of the object we wish to scan here *viz.* the fluorophore is a sphere. Such spheres serve as good models for fluorescent-coated beads.

These spheres can be generated from their analytical expressions in the frequency domain as (*cf.* Hanser, et al. [2003]; Lai, et al. [2005])

$$O(\mathbf{k}) = \frac{\sin(2\pi R\mathbf{k}_\rho) - (2\pi R\mathbf{k}_\rho)\cos(2\pi R\mathbf{k}_\rho)}{\pi^2(\mathbf{k}_\rho)^3} \tag{A.1}$$

where R is the radius of the desired sphere, and the sampling in the frequency domain is carried out so that $\Delta_{k_x} = 1/N_x$, $\Delta_{k_y} = 1/N_y$, and $\Delta_{k_z} = \Delta_{xy}/(\Delta_z N_z)$. Δ_{xy} and Δ_z are the radial and the axial samplings in the image space, and $\mathbf{k}_\rho = (k_x^2 + k_y^2 + k_z^2)^{1/2}$. To avoid Gibbs phenomenon, the numerical approximation of $O(\mathbf{k})$ was multiplied by a ℓ_∞ normalized 3D Gaussian function with variances $(N_x, N_y, N_z)/(2\pi)$.

APPENDIX B

Maximum Likelihood Expectation Maximization

Given the observation $i(\mathbf{x})$ and the PSF $h(\mathbf{x})$, when the requirement is only to find the original object $o(\mathbf{x})$, we use the Bayes' theorem

$$\Pr(o(\mathbf{x})|i(\mathbf{x}')) = \frac{\Pr(i(\mathbf{x}')|o(\mathbf{x}))\Pr(o(\mathbf{x}))}{\sum\limits_{\mathbf{x}''\in\Omega_s}\Pr(i(\mathbf{x})|o(\mathbf{x}''))\Pr(o(\mathbf{x}''))}; \mathbf{x},\mathbf{x}',\mathbf{x}''\in\Omega_s. \tag{B.1}$$

Considering all the $i(\mathbf{x}')$ and its dependance on $o(\mathbf{x})$, we can say

$$\Pr(o(\mathbf{x})) = \sum_{\mathbf{x}'\in\Omega_s}\Pr(o(\mathbf{x})i(\mathbf{x}')) \tag{B.2}$$

$$= \sum_{\mathbf{x}'\in\Omega_s}\Pr(o(\mathbf{x})|i(\mathbf{x}'))\Pr(i(\mathbf{x}')), \tag{B.3}$$

since $\Pr(o(\mathbf{x})|i(\mathbf{x}')) = \Pr(o(\mathbf{x})i(\mathbf{x}'))/\Pr(i(\mathbf{x}'))$. Substituting Eq. B.1 in Eq. B.3 we can say

$$\Pr(o(\mathbf{x})) = \sum_{\mathbf{x}'\in\Omega_s}\frac{\Pr(i(\mathbf{x}')|o(\mathbf{x}))\Pr(o(\mathbf{x}))\Pr(i(\mathbf{x}'))}{\sum\limits_{\mathbf{x}''\in\Omega_s}\Pr(i(\mathbf{x})|o(\mathbf{x}''))\Pr(o(\mathbf{x}''))} \tag{B.4}$$

We see that the left and the right hand sides of the above equation have both the desired result, $\Pr(o(\mathbf{x}))$. Hence, we write the result in the following iterative form:

$$\Pr^{(n+1)}(o(\mathbf{x})) = \Pr^{(n)}(o(\mathbf{x}))\sum_{\mathbf{x}'\in\Omega_s}\frac{\Pr(i(\mathbf{x}')|o(\mathbf{x}))\Pr(i(\mathbf{x}'))}{\sum\limits_{\mathbf{x}''\in\Omega_s}\Pr(i(\mathbf{x}')|o(\mathbf{x}''))\Pr^{(n)}(o(\mathbf{x}''))}, \tag{B.5}$$

where $n = \{0, 1, \ldots\}$, and the initial $\Pr^{(0)}(o(\mathbf{x}))$ is assumed to be known or estimated. Using Bayes' postulate, a uniform distribution is assumed for the initial estimate so that

$$\Pr^{(0)}(o(\mathbf{x})) = \frac{\hat{o}^{(0)}(\mathbf{x})}{\sum_{\mathbf{x}\in\Omega_s} o(x)}, \tag{B.6}$$

and

$$\Pr^{(n)}(o(\mathbf{x})) = \frac{\hat{o}^{(n)}(\mathbf{x})}{\sum_{\mathbf{x}\in\Omega_s} o(x)}. \tag{B.7}$$

By the property of flux conservation, $\sum_{\mathbf{x}\in\Omega_s} i(x) = \sum_{\mathbf{x}\in\Omega_s} o(x)$, hence

$$\Pr(i(\mathbf{x})) = \frac{i(\mathbf{x})}{\sum_{\mathbf{x}\in\Omega_s} i(x)} \tag{B.8}$$

$$= \frac{i(\mathbf{x})}{\sum_{\mathbf{x}\in\Omega_s} o(x)}. \tag{B.9}$$

Similarly,

$$\Pr(i(\mathbf{x}')|o(\mathbf{x})) = \Pr(h(\mathbf{x}' - \mathbf{x})) \tag{B.10}$$

$$= \frac{h((\mathbf{x}' - \mathbf{x}))}{\sum_{\mathbf{x}''\in\Omega_s} h(\mathbf{x}'')}. \tag{B.11}$$

So in an iterative form, Eq. B.5 becomes

$$\left(\frac{\hat{o}^{(n+1)}(\mathbf{x})}{\sum\limits_{\mathbf{x}\in\Omega_s} \hat{o}^{(n)}(\mathbf{x})}\right) = \left(\frac{\hat{o}^{(n)}(\mathbf{x})}{\sum\limits_{\mathbf{x}\in\Omega_s} o(\mathbf{x})}\right) \sum_{\mathbf{x}'\in\Omega_s} \frac{\left(\frac{h(\mathbf{x}'-\mathbf{x})}{\sum\limits_{\mathbf{x}''\in\Omega_s} h(\mathbf{x}'')}\right)\left(\frac{i(\mathbf{x}')}{\sum\limits_{\mathbf{x}\in\Omega_s} o(\mathbf{x})}\right)}{\sum\limits_{\mathbf{x}''\in\Omega_s}\left(\frac{h(\mathbf{x}'-\mathbf{x}'')}{\sum\limits_{\mathbf{x}''\in\Omega_s} h(\mathbf{x}'')}\right)\left(\frac{\hat{o}^{(n)}(\mathbf{x}'')}{\sum\limits_{\mathbf{x}\in\Omega_s} o(\mathbf{x})}\right)} \tag{B.12}$$

or

$$\hat{o}^{(n+1)}(\mathbf{x}) = \hat{o}^{(n)}(\mathbf{x}) \sum_{\mathbf{x}'\in\Omega_s} \frac{h(\mathbf{x}' - \mathbf{x})i(\mathbf{x}')}{\sum\limits_{\mathbf{x}''\in\Omega_s} h(\mathbf{x}' - \mathbf{x}'')\hat{o}^{(n)}(\mathbf{x}'')} \tag{B.13}$$

Appendix C

Fourier Transform of a Gaussian

Definition 9. A function of three independent variables is called *separable* with respect to a specific co-ordinate system if it can be written as a product of three functions, each of which depends only on one of the independent variables. Thus, the function f is separable in co-ordinates (x, y, z) if

$$f(x,y,z) = f_X(x) f_Y(y) f_Z(z). \tag{C.1}$$

The 3D Gaussian function, $h : \mathbb{N}^3 \to \mathbb{R}_+$, is considered to be separable as the function can be written individually as the combination of three 1D Gaussian functions. The 3D convolution of h with the function $o : \mathbb{N}^3 \to \mathbb{R}_+$ is thus reduced to three successive 1D multiplications in the Fourier domain with the Fourier transform of o. Thus,

$$(h * o)(x,y,z) = (h(x) * (h(y) * (h(z) * o(x,y,z)))), \tag{C.2}$$

and

$$(h*o)(x,y,z) = \mathscr{F}_{3D}^{-1} \times (\mathscr{F}_{1D}(h(z)) \times (\mathscr{F}_{1D}(h(y)) \times (\mathscr{F}_{1D}(h(x)) \times \mathscr{F}_{3D}(o(x,y,z))))). \tag{C.3}$$

The Fourier transform of a continuous function $h(\mathbf{x})$ is given by

$$\mathscr{F}(h(x)) = H(k_x) = \int_{-\infty}^{+\infty} h(x) e^{-j2\pi k_x x}\, dx\,, \tag{C.4}$$

where $j^2 = -1$, $h(x) = (1/(2\pi)^{1/2}\sigma_x)\exp(-x^2/(2\sigma_x^2))$ is the 1D Gaussian function, and k_x is the co-ordinate in the frequency domain. A closed-form expression for $H(k_x)$ exists and the analytical expression can be written as

$$\begin{aligned} H(k_x) &= \int_{-\infty}^{+\infty} \frac{1}{((2\pi)^{\frac{1}{2}}\sigma_x)} \exp\left(-\frac{x^2}{2\sigma_x^2}\right) \exp(-j2\pi k_x x)\, dx \\ &= \exp\left(-\frac{(2\pi k_x \sigma_x)^2}{2}\right). \end{aligned} \tag{C.5}$$

From the above expression, it is clear that the Fourier transform of a Gaussian is also a Gaussian. The proof is not straight forward and it is as below.

Proof. It is simple to show that differentiation of $h(x)$ with respect to x is

$$\frac{d}{dx} h(x) = \left(\frac{-x}{\sigma_x^2}\right) h(x). \tag{C.6}$$

Applying the Fourier transform to the above equation and using the differentiation property of the Fourier transform, we get

$$\begin{aligned} (j2\pi k_x) H(k_x) &= \frac{-1}{\sigma_x^2} \int_{-\infty}^{+\infty} x h(x) \exp(-j2\pi k_x x)\, dx\,, \\ &= \frac{-j}{\sigma_x^2} \frac{d}{dk_x} H(k_x)\,. \end{aligned} \tag{C.7}$$

The above expression can be simplified to

$$\frac{1}{H(k_x)} \frac{d}{dk_x} H(k_x) = -2\pi k_x \sigma_x^2. \tag{C.8}$$

Integrating both sides, we get

$$\int_0^{k_x} \frac{1}{H(k_x')} \frac{d}{dk_x'} H(k_x')\, dk_x' = -2\pi\sigma_x^2 \int_0^{k_x} k_x'\, dk_x'\,, \tag{C.9}$$

or equivalently

$$\log(H(k_x)) - \log(H(0)) = -\frac{\sigma_x^2}{2}(2\pi k_x)^2. \tag{C.10}$$

Thus, we arrive at the following expression:

$$H(k_x) = H(0)\exp\left(-\frac{(2\pi\sigma_x k_x)^2}{2}\right), \tag{C.11}$$

with $H(0) = 1$. □

APPENDIX D

Numerical Implementation for Total Variation Regularization

The deconvolution algorithm introduced in Ch. 4.2.1 uses the TV regularization as a constraint on the object. In this appendix, we provide the numerical approximation of the TV regularization function used in the iterative algorithm.

In the continuous domain, the 3D TV semi-norm is written as

$$\mathrm{TV}(o(\mathbf{x})) = \int\limits_{\mathbf{x}\in\mathbb{R}^3} |\nabla o(\mathbf{x})| d\mathbf{x} \tag{D.1}$$

$$= \iiint\limits_{(x,y,z)\in\mathbb{R}^3} \left(\left(\frac{\partial o(\mathbf{x})}{\partial x}\right)^2 + \left(\frac{\partial o(\mathbf{x})}{\partial y}\right)^2 + \left(\frac{\partial o(\mathbf{x})}{\partial z^2}\right)^2 \right)^{\frac{1}{2}} dx\, dy\, dz\,. \tag{D.2}$$

We recall the discrete smooth form of the above equation that was written in Eq. (2.23) as

$$\begin{aligned} \mathscr{J}_{\mathrm{reg},o}(o(\mathbf{x})) &= \lambda_o \sum_{\mathbf{x}\in\Omega_s} |\nabla o(\mathbf{x})|_\epsilon \\ &= \lambda_o \sum_{\mathbf{x}\in\Omega_s} \left(\epsilon^2 + |\nabla o(\mathbf{x})|^2\right)^{\frac{1}{2}}. \end{aligned} \tag{D.3}$$

An arbitrary parameter $\epsilon < 10^{-3}$ is introduced to ensure that the norm $|\nabla o(\mathbf{x})|$ is differentiable. In order to estimate the object, by using Eq. (4.9), it is necessary to calculate the discrete form of the term $\mathrm{div}(\nabla o(\mathbf{x})/|\nabla o(\mathbf{x})|)$.

Definition 10. We define the *minmod* function $m(a,b)$ as

$$m(a,b) \overset{\mathrm{def}}{=} \left(\frac{\mathrm{sign}(a)+\mathrm{sign}(b)}{2}\right) \min(|a|,|b|), \tag{D.4}$$

where the function sign(a) is

$$\operatorname{sign}(a) = \begin{cases} 1, & a > 0, \\ -1, & a < 0, \\ 0, & a = 0 \end{cases} \tag{D.5}$$

A stable numerical scheme is the following (*cf.* Rudin, et al. [1992]; Dey, et al. [2004]):

$$\begin{aligned} \operatorname{div}\left(\frac{\nabla o(\mathbf{x})}{|\nabla o(\mathbf{x})|}\right) = \quad & \Bigg(\Delta_-^x \frac{\Delta_+^x o_{ijk}}{\left(\epsilon^2 + (\Delta_+^x o_{ijk})^2 + m(\Delta_+^y o_{ijk}, \Delta_-^y o_{ijk})^2 + m(\Delta_+^z o_{ijk}, \Delta_-^z o_{ijk})^2\right)^{\frac{1}{2}}} \\ & + \frac{h_x}{h_y}\Delta_-^y \frac{\Delta_+^y o_{ijk}}{\left(\epsilon^2 + (\Delta_+^y o_{ijk})^2 + m(\Delta_+^z o_{ijk}, \Delta_-^z o_{ijk})^2 + m(\Delta_+^x o_{ijk}, \Delta_-^x o_{ijk})^2\right)^{\frac{1}{2}}} \\ & + \frac{h_x}{h_z}\Delta_-^z \frac{\Delta_+^z o_{ijk}}{\left(\epsilon^2 + (\Delta_+^z o_{ijk})^2 + m(\Delta_+^x o_{ijk}, \Delta_-^x o_{ijk})^2 + m(\Delta_+^y o_{ijk}, \Delta_-^y o_{ijk})^2\right)^{\frac{1}{2}}}\Bigg)\frac{1}{h_x}, \end{aligned} \tag{D.6}$$

where the numerical forward and backward derivatives are defined as

$$\begin{aligned} \Delta_\pm^x o_{ijk} &= \frac{\pm(o_{(i\pm1)jk} - o_{ijk})}{h_x}, \\ \Delta_\pm^y o_{ijk} &= \frac{\pm(o_{i(j\pm1)k} - o_{ijk})}{h_y}, \\ \Delta_\pm^z o_{ijk} &= \frac{\pm(o_{ij(k\pm1)} - o_{ijk})}{h_z}. \end{aligned} \tag{D.7}$$

$(i,j,k) \in \mathbb{Z}$ are the indices on a 3D grid $[0, N_x - 1] \times [0, N_y - 1] \times [0, N_z - 1]$, and $(i\Delta_{xy}, j\Delta_{xy}, k\Delta_z)$ are their distances from the central plane. As mentioned earlier, $(\Delta_{xy}, \Delta_{xy}, \Delta_z)$ are the Nyquist sampled voxel sizes. The advantage of using the minmod function, $m(a,b)$, is that it equals zero iff a and b, in Eq. (D.4), have the opposite sign. Thus, if $\Delta_+^{x,y,z} o_{ijk}$ and $\Delta_-^{x,y,z} o_{ijk}$ do not have the same sign, we choose $\Delta_\pm^{x,y,z} o_{ijk} = 0$.

In Eqs. (D.7) and (D.7), (h_x, h_y, h_z) are chosen as $h_x = 1$, $h_y = 1$ and $h_z = \Delta_{xy}/\Delta_z$. In the boundaries of the image, we use the following limits (in the continuous domain):

$$\frac{\partial}{\partial \vec{n}} o(x) = 0, \tag{D.8}$$

where $\vec{n}$ is normal to the boundary $\partial \Omega_s$ of Ω_s. For a discrete functional, this leads to the following symmetric boundary conditions:

$$\begin{aligned} o_{0jk} &= o_{1jk}, & o_{i0k} &= o_{i1k}, & o_{ij0} &= o_{ij1} \\ o_{(N_x+1)jk} &= o_{(N_x)jk}, & o_{i(N_y+1)k} &= o_{i(N_y)k}, & o_{ij(N_z+1)} &= o_{ij(N_z)}. \end{aligned} \tag{D.9}$$

List of Publications and Scientific Activities of the Author

Journal Papers

1. P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin and J. Zerubia. *Blind deconvolution for thin-layered confocal imaging*. Appl. Opt., vol. 48, no. 22, pages 4437–4448, 2009.

International Conferences

1. L. Blanc-Féraud, P. Pankajakshan, M. Carlavan, J. Zerubia and J.-C. Olivo-Marin. *Methods for Confocal Microscpic 3D Image Restoration*. SIAM Conference on Imaging Science, accepted, Chicago, USA, April 2010.
2. P. Pankajakshan, L. Blanc-Féraud, Z. Kam and J. Zerubia. *Point-spread function retrieval in fluorescence microscopy*. Proc. IEEE International Symposium on Biomedical Imaging, pages 1095–1098, Boston, USA, July 2009.
3. P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin and J. Zerubia. *Blind deconvolution for diffraction-limited fluorescence microscopy*. Proc. IEEE International Symposium on Biomedical Imaging, pages 740–743, Paris, France, May 2008.
4. P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin and J. Zerubia. *Parametric blind deconvolution for confocal laser scanning microscopy*. Proc. IEEE Eng. Med. Biol. Soc., pages 6531–6534, Lyon, France, August 2007.

List of Research Reports

1. P. Pankajakshan, L. Blanc-Féraud, Z. Kam, and J. Zerubia. *Space non-invariant point-spread function and its estimation in fluorescence microscopy*. Research Report 7157, INRIA, France, December 2009.
2. P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin and J. Zerubia. *Parametric blind deconvolution for confocal laser scanning microscopy (CLSM)-proof of concept*. Research Report 6493, INRIA, France, March 2008.

Invited Talks

1. '*On blind deconvolution and point-spread function retrieval,*' Quantitative Image Analysis Unit, Pasteur Institute, Paris, France, June 19, 2009.
2. '*On deconvolution and the hidden "phase" in an incoherent imaging system,*' Cross Seminar, INRIA, Sophia-Antipolis, France, April 07, 2009.
3. '*Restoration of confocal laser scanning microscopic images,*' CNES, Toulouse, France, January 14, 2009.
4. '*Myopic deconvolution for fluorescence microscopy,*' LabEl, Laboratory MIPS, Université de Haute-Alsace, Mulhouse, France, December 18, 2008.
5. '*Blind deconvolution and 3D PSF modeling in biological microscopy,*' High council for research, scientific and technological co-operation, status seminar program in medical imaging, Paris, France, November 17, 2008.
6. '*Inverse problems in image processing,*' Department of Mathematics, Indian Institute of Science (IISc) Math Initiative (IMI), Bangalore, India, August 06, 2008.
7. '*Inverse problems in microscopy,*' Department of Electrical Engineering, Indian Institute of Technology (IIT) Roorkee, India, September 15, 2006.

Bibliography

D. A. Agard (1984). 'Optical sectioning microscopy: cellular architecture in three dimensions'. *Ann. Rev. Biophys. Bioeng.* **13**:191–219. 10, 22, 23, 27, 28

S. Alenius & U. Ruotsalainen (1997). 'Bayesian image reconstruction for emission tomography based on median root prior'. *Eur. J. Nucl. Med. Mol. Imaging* **24**(3):258–265. 33, 34

K. E. Atkinson (1989). *An Introduction to Numerical Analysis*. Wiley, New York, 2nd edn. 82

G. Aubert & P. Kornprobst (2001). *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, vol. 147 of *Applied Mathematical Sciences*. Springer Verlag. 36

G. R. Ayers & J. C. Dainty (1988). 'Iterative blind deconvolution method and its applications'. *Opt. Lett.* **13**(7):547–549. 42, 92, 93

J. Bardsley, et al. (2006). 'Blind iterative restoration of images with spatially-varying blur'. *Optics Express* **14**:1767–1782. 137

J. C. Bezdek, et al. (1987). 'Local convergence analysis of a grouped variable version of coordinate descent'. *J. Optim. Theory Appl.* **54**(3):471–477. 38

D. S. C. Biggs & M. Andrews (1997). 'Acceleration of iterative image restoration algorithms'. *Appl. Opt.* **36**(8):1766–1775. 41, 59, 104

T. E. Bishop, et al. (2007). 'Blind image deconvolution: problem formulation and existing approaches'. In P. Campisi & K. Egiazarian (eds.), *Blind Image Deconvolution: Theory and Applications*, chap. 1, pp. 1–42. CRC Press. 42

L. Blanc-Féraud, et al. (2009). 'Déconvolution aveugle d'image'. In A. Djafari (ed.), *Problèmes Inverses en Imagerie et en Vision*, Traité IC2, Traitement du Signal et de l'Image, pp. 107–132. Hermès Publ. 39

M. J. Booth (2007). 'Adaptive optics in microscopy'. *Philos. Transact. A Math. Phys. Eng. Sci.* **365**(1861):2829–2843. 110

M. J. Booth, et al. (2002). 'Adaptive aberration correction in a confocal microscope'. *Proc. Natl. Acad. Sci.* **99**(9):5788–5792. 110

M. J. Booth, et al. (1998). 'Aberration correction for confocal imaging in refractive-index-mismatched media'. *J. Microsc.* **192**:90–98. 66, 108, 133

M. Born & E. Wolf (1999). *Principles of Optics*. Cambridge U. Press. 3, 6, 48, 52, 53, 65, 135

S. P. Boyd & L. Vandenberghe (2004). *Convex Optimization*. Cambridge University Press. 29, 36

E. Bratsolis & M. Sigelle (2001). 'A spatial regularization method preserving local photometry for Richardson-Lucy restoration'. *Astron. Astrophys.* **375**:1120–1128. 80

L. M. Bregman (1965). 'The method of successive projection for finding a common point of convex sets'. *Soviet Math. Dokl.* **6**:688–692. 27, 106

C. Brune, et al. (2009). 'Bregman-EM-TV methods with application to optical nanoscopy'. In *Proc. Second International Conference on Scale Space and Variational Methods in Computer Vision*, vol. 5567 of *Lecture Notes in Computer Science*, pp. 235–246, Berlin, Heidelberg. Springer-Verlag. 105

A. Buades, et al. (2005). 'A review of image denoising algorithms, with a new one'. *Multiscale Model. Simul.* **4**(2):490–530. 30

P. Campisi & K. Egiazarian (eds.) (2007). *Blind Image Deconvolution: Theory and Applications*. CRC Press. 42

M. B. Cannell, et al. (2006). 'Image enhancement by deconvolution'. In J. B. Pawley (ed.), *Handbook of Biological Confocal Microscopy*, chap. 25, pp. 488–500. Springer, 3rd edn. 22

A. S. Carasso (1999). 'Linear and nonlinear image deblurring: a documented study'. *SIAM J. Numer. Anal.* **36**(6):1659–1689. 22

A. S. Carasso (2001). 'Direct blind deconvolution'. *SIAM J. Appl. Math* **61**(6):1980–2007. 42

A. S. Carasso (2003). 'The APEX method in image sharpening and the use of low exponent lévy stable laws'. *SIAM J. Appl. Math.* **63**(2):593–618. 42

M. Carlavan, et al. (2009). 'Complex wavelet regularization for solving inverse problems in remote sensing'. In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, Cape Town, South Africa. 34

W. A. Carrington, et al. (1990). *3D Fluorescence Imaging of Single Cells Using Image Restoration*. Wiley-Liss. 23, 28

T. F. Chan & J. Shen (2005). *Image Processing and Analysis: Variational, PDE, Wavelet, and Stochastic Methods*. SIAM Publisher. 39

T. F. Chan & C. K. Wong (1998). 'Total variation blind deconvolution'. *IEEE Trans. Image Process.* **7**(3):370–375. 59

P. Charbonnier, et al. (1997). 'Deterministic edge-preserving regularization in computed imaging'. *IEEE Trans. Image Process.* **6**(2):298–311. 34, 36

P. Charbonnier, et al. (1993). 'An adaptive reconstruction method involving discontinuities'. In *IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 5, pp. 491–494, Minneapolis, MN, USA. 34

C. Chaux, et al. (2007). 'Wavelet-based restoration methods: application to 3D confocal microscopy images'. In *Proc. SPIE*, vol. 6701, San Diego, USA. 30, 37

T.-W. Chen, et al. (2006). 'In situ background estimation in quantitative fluorescence imaging'. *Biophys. J.* **90**(7):2534–2547. 12

G. Chenegros, et al. (2006). '3-D deconvolution of adaptive-optics corrected retinal images'. In J.-Á. Conchello, C. J. Cogswell, & T. Wilson (eds.), *Proc. SPIE*, vol. 6090, pp. 144–151, San Jose, CA, USA. 34

J.-Á. Conchello & J. G. McNally (1996). 'Fast regularization technique for expectation maximization algorithm for computational optical sectioning microscopy'. In C. J. Cogswell, G. S. Kino, & T. Wilson (eds.), *Three-Dimensional microscopy: image acquisition and processing*, vol. 2655, pp. 199–208. SPIE. 32, 34

I. Csiszar (1991). ‘Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems’. *Ann. Stat.* **19**(4):2033–2066. 18

J. C. Dainty & M. A. Fiddy (1984). ‘The essential role of prior knowledge in phase retrieval’. *J. Mod. Opt.* **31**(3):325–330. 111

J. B. de Monvel, et al. (2001). ‘Image restoration for confocal microscopy: improving the limits of deconvolution, with application to the visualization of the mammalian hearing organ’. *Biophys. J.* **80**(5):2455–2470. 30, 40

M. de Moraes Marim, et al. (2008). ‘Improving single particle localization with an empirically calibrated Gaussian kernel’. In *Proc. IEEE International Symposium on Biomedical Imaging*, pp. 1003–1006. 101

G. Demoment (1989). ‘Image reconstruction and restoration: overview of common estimation structures and problems’. *IEEE Trans. Acoust. Speech Signal Process.* **37**(12):2024–2036. 31

N. Dey, et al. (2006). ‘Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution’. *Microsc. Res. Tech.* **69**:260–266. 23, 32, 34, 36, 78

N. Dey, et al. (2004). ‘3D microscopy deconvolution using Richardson-Lucy algorithm with total variation regularization’. Research Report 5272, INRIA, Sophia-Antipolis, France. 7, 29, 36, 99, 152

A. Dieterlen, et al. (2004). ‘Identification and restoration in 3D fluorescence microscopy’. In O. V. Angelsky (ed.), *Proc. SPIE*, vol. 5477, pp. 105–113. SPIE. 30, 67

H. Du, et al. (1998). ‘PhotochemCAD: A computer-aided design and research tool in photochemistry’. *Photochemistry and Photobiology* **68**(2):141–142. 141

F. X. Dupé, et al. (2008). ‘Image deconvolution under Poisson noise using sparse representations and proximal thresholding iteration’. In *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, pp. 761–764. 30, 31, 34

A. Egner, et al. (1998). 'Refractive index mismatch induced intensity and phase variations in fluorescence confocal, multiphoton and 4Pi-microscopy'. *Opt. Comm.* **153**:211–217. 108

A. Erhardt, et al. (1985). 'Reconstructing 3-D light-microscopic images by digital image processing'. *Appl. Opt.* **24**:194–200. 23, 25, 26

R. Fergus, et al. (2006). 'Removing camera shake from a single photograph'. *Proc. ACM Trans. on Graphics* **25**:787–794. 33

M. A. T. Figueiredo & R. D. Nowak (2001). 'Wavelet-based image estimation: an empirical bayes approach using Jeffreys' noninformative prior'. *IEEE Trans. Image Process.* **10**(9):1322–1331. 37

M. A. T. Figueiredo & R. D. Nowak (2003). 'An EM algorithm for wavelet-based image restoration'. *IEEE Trans. Image Process.* **12**(8):906–916. 23, 32, 34

R. W. Gerchberg & W. O. Saxton (1972). 'A practical algorithm for the determination of the phase from image and diffraction plane pictures'. *Optik* **35**:237–246. 42

S. F. Gibson & F. Lanni (1989). 'Diffraction by a circular aperture as a model for three-dimensional optical microscopy'. *J. Opt. Soc. Am. A* **A6**:1357–1367. 54, 65, 76, 126

R. C. Gonzalez & R. E. Woods (1987). *Digital image processing*. Addison-Wesley Reading, Mass. 10, 26

J. W. Goodman (2004). *Introduction to Fourier Optics*. Roberts & Company Publishers. 10

B. M. Hanser, et al. (2003). 'Phase retrieval for high-numerical-aperture optical systems'. *Opt. Lett.* **28**(10):801–803. 55, 109, 143

B. M. Hanser, et al. (2001). 'Phase retrieval of widefield microscopy point spread functions'. In J.-A. Conchello, C. J. Cogswell, & T. Wilson (eds.), *Proc. SPIE*, vol. 4261, pp. 60–68. 109

G. Harikumar & Y. Bresler (1999). 'Perfect blind restoration of images blurred by multiple filters: theory and efficient algorithms'. *IEEE Trans. Image Process.* **8**:202–219. 141

S. W. Hell & E. H. K. Stelzer (2006). 'Lens aberrations in confocal fluorescence microscopy'. In J. B. Pawley (ed.), *Handbook of Biological Confocal Microscopy*, chap. 11. Springer, 3 edn. 65

Y. Hiraoka, et al. (1990). 'Determination of three-dimensional imaging properties of a light microscope system'. *Biophys. J.* **57**:325–333. 47, 110, 126

T. J. Holmes (1988). 'Maximum-likelihood image restoration adapted for noncoherent optical imaging'. *J. Opt. Soc. Am. A* **5**:666–673. 23

T. J. Holmes (1992). 'Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach'. *J. Opt. Soc. Am. A* **9**(7):1052–1061. 41, 42, 59, 91, 93, 104

E. F. Y. Hom, et al. (2007). 'AIDA: an adaptive image deconvolution algorithm with application to multi-frame and three-dimensional data'. *J. Opt. Soc. Am. A* **24**(6):1580–1600. 32, 34, 42

H. H. Hopkins (1955). 'The frequency response of a defocused optical system'. *Proc. R. Soc. London Ser. A.* **231**:91–103. 56, 57

R. Hudson, et al. (1996). 'Optical microscopy system for 3D dynamic imaging'. In C. J. Cogswell, G. S. Kino, & T. Wilson (eds.), *Proc. SPIE*, vol. 2655, pp. 187–198. 10

S. Inoué (2006). 'Foundations of confocal scanned imaging in light microscopy'. In J. B. Pawley (ed.), *Handbook of Biological Confocal Microscopy*, chap. 1, pp. 1–19. Springer, 3rd edn. 2

A. Jalobeanu, et al. (2002). 'Hyperparameter estimation for satellite image restoration using a MCMC maximum likelihood method'. *Patt. Recog.* **35**(2):341–352. 85

A. Jalobeanu, et al. (2003). 'Satellite image debluring using complex wavelet packets'. *Int. J. Comput. Vis.* **51**(3):205–217. 37

A. Jalobeanu, et al. (2007). 'Bayesian estimation of blur and noise in remote sensing imaging'. In P. Campisi & K. Egiazarian (eds.), *Blind Image Deconvolution: Theory and Applications*, chap. 6, pp. 239–275. CRC Press. 39

M. Jiang & G. Wang (2003). 'Development of blind image deconvolution and its applications'. *J. X-Ray Sci. Technol.* **11**:13–19. 18

S. Joshi & M. I. Miller (1993). 'Maximum a posteriori estimation with Good's roughness for three-dimensional optical-sectioning microscopy'. *J. Opt. Soc. Am. A* **10**(5):1078. 23

R. Juskaitis (2006). 'Measuring the real point spread function of high numerical aperture microscope objective lenses'. In J. B. Pawley (ed.), *Handbook of Biological Confocal Microscopy*, chap. 11, pp. 239–250. Springer, 3$^{\text{rd}}$ edn. 65, 110

Z. Kam, et al. (1997). 'Three-dimensional microscopy in thick biological samples: A fresh approach for adjusting focus and correcting spherical aberration'. *Bioimaging* **5**(1):40–49. 110

Z. Kam, et al. (2001). 'Computational adaptive optics for live three-dimensional biological imaging'. *Proc. Natl. Acad. Sci.* **98**(7):3790–3795. 110, 138

Z. Kam, et al. (2007). 'Modelling the application of adaptive optics to wide-field microscope live imaging'. *J. Microsc.* **226**:33–42. 110, 134, 136

C. Kervrann & A. Trubuil (2004). 'An adaptive window approach for Poisson noise reduction and structure preserving in confocal microscopy'. In *Proc. IEEE International Symposium on Biomedical Imaging*, vol. 1, pp. 788–791, Arlington, VA, USA. 30

M. Koshy, et al. (1990). 'Solution of toeplitz systems for the restoration of 3-D optical sectioning microscopy data'. In L. C. Smith (ed.), *Proc. SPIE*, vol. 1205, pp. 64–71. 27

D. Kundur & D. Hatzinakos (1996). 'Blind image restoration via recursive filtering using deterministic constraints'. In *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 4, pp. 2283–2286, Atlanta, Georgia. 42

D. Kundur & D. Hatzinakos (1998). 'A novel blind deconvolution scheme for image restoration using recursive filtering'. *IEEE Trans. Signal Process.* **46**(2):375–390. 42

X. Lai, et al. (2005). 'Noise suppression of point spread functions and its influence on deconvolution of three-dimensional fluorescence microscopy image sets'. *J. Microsc.* **217**(1):93–108. 107, 143

A. Levin, et al. (2009). 'Understanding and evaluating blind deconvolution algorithms'. In *Proc. IEEE Comp. Vis. and Pat. Recog.*, Miami, FL, USA. to appear. 40, 42

L. B. Lucy (1974). 'An iterative technique for the rectification of observed distributions'. *Astron. J.* **79**:745–754. 29

F. Luisier, et al. (2010). 'Fast interscale wavelet denoising of Poisson-corrupted images'. *Signal Process.* **90**(2):415 – 427. 30

E. Maalouf, et al. (2008). 'Fast deconvolution with non-invariant PSF for 3-D fluorescence microscopy'. In P. Schelkens, T. Ebrahimi, G. Cristóbal, & F. Truchetet (eds.), *Proc. SPIE*, vol. 7000, pp. 70001K–70001K–8, Strasbourg, France. SPIE. 137

L. Mandel (1979). 'Sub-Poissonian photon statistics in resonance fluorescence'. *Opt. Lett.* **4**:205–207. 9

J. Marc, et al. (1998). 'A GFP-MAP4 reporter gene for visualizing cortical microtubule rearrangements in living epidermal cells'. *The Plant Cell* **10**(11):1927–1939. 102

J. Markham & J.-A. Conchello (1999). 'Parametric blind deconvolution: a robust method for the simultaneous estimation of image and blur'. *J. Opt. Soc. Am. A* **16**(10):2377–2391. 42

J. G. McNally, et al. (1994). 'Artifacts in computational optical-sectioning microscopy'. *J. Opt. Soc. Am. A* **11**:1056–1067. 47, 76

E. S. Meinel (1986). 'Origins of linear and nonlinear recursive restoration algorithms'. *J. Opt. Soc. Am. A* **3**(6):787–799. 22

O. V. Michailovich & D. R. Adam (2007). ‘Deconvolution of medical images from microscopic to whole body images’. In P. Campisi & K. Egiazarian (eds.), *Blind Image Deconvolution: Theory and Applications*, chap. 5, pp. 169–237. CRC Press. 42

K. Miller (1970). ‘Least squares methods for ill-posed problems with a prescribed bound’. *SIAM J. Math. Anal.* **1**(1):52–74. 34, 36

M. Minsky (1988). ‘Memoir on inventing the confocal scanning microscope’. *Scanning* **10**:128–138. 2

A. Mohammad-Djafari (1996). ‘A full Bayesian approach for inverse problems’. In K. Hanson & R. N. Silver (eds.), *Maximum entropy and Bayesian methods*, vol. 79, pp. 135–143. Kluwer Academic. 85

J. R. Monck, et al. (1992). ‘Thin-section ratiometric Ca2+ images obtained by optical sectioning of fura-2 loaded mast cells’. *J. Cell Biol.* **116**(3):745–759. 22, 23

Y. Nesterov (2004). *Introductory Lectures on Convex Optimization: A Basic Course*. Kluwer Academic. 36

J. Nocedal & S. J. Wright (1999). ‘Quasi-Newton methods’. In *Numerical Optimization*, chap. 8, pp. 193–201. Springer-Verlag, New York. 114, 115, 119

P. Pankajakshan, et al. (2009a). ‘Point-spread function retrieval in fluorescence microscopy’. In *Proc. IEEE International Symposium on Biomedical Imaging*, pp. 1095–1098, Boston, USA. 110

P. Pankajakshan, et al. (2007). ‘Parametric blind deconvolution for confocal laser scanning microscopy’. In *Proc. IEEE Eng. Med. Biol. Soc.*, pp. 6531–6534, Lyon, France. 32

P. Pankajakshan, et al. (2008a). ‘Blind deconvolution for diffraction-limited fluorescence microscopy’. In *Proc. IEEE International Symposium on Biomedical Imaging*, pp. 740–743, Paris, France. 101

P. Pankajakshan, et al. (2008b). ‘Parametric blind deconvolution for confocal laser scanning microscopy (CLSM)-proof of concept’. Research Report 6493, INRIA, Sophia-Antipolis, France. 22, 27, 30, 34, 36

P. Pankajakshan, et al. (2009b). 'Blind deconvolution for thin-layered confocal imaging'. *Appl. Opt.* **48**(22):4437–4448. 32, 86, 101

E. Pantin, et al. (2007). 'Deconvolution and blind deconvolution in astronomy'. In P. Campisi & K. Egiazarian (eds.), *Blind Image Deconvolution: Theory and Applications*. CRC Press. 34, 36

J. B. Pawley (1994). *Three-Dimensional Confocal Microscopy: Volume Investigation of Biological Specimens*, chap. Sources of noise in three-dimensional microscopical data sets, pp. 47–94. Academic Press. 7

J. B. Pawley (2006). 'Fundamental limits in confocal microscopy'. In J. B. Pawley (ed.), *Handbook of Biological Confocal Microscopy*, chap. 2, pp. 20–42. Springer, 3rd edn. 3, 40, 108

T. Peng, et al. (2007). 'Algorithms for on-line monitoring of micro spheres in an optical tweezers-based assembly cell'. *Journal of Computing and Information Science in Engineering* **7**(4):330–338. 121

M. Persson, et al. (2001). 'Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography'. *Phys. Med. Bio.* **46**(3):853–866. 34, 36

C. Preza & J.-A. Conchello (2004). 'Depth-variant maximum-likelihood restoration for three-dimensional fluorescence microscopy'. *J. Opt. Soc. Am. A* **21**(9):1593–1601. 136, 137, 140, 141

C. Preza, et al. (1992). 'Regularized linear method for reconstruction of three-dimensional microscopic objects from optical sections'. *J. Opt. Soc. Am. A* **9**(2):219–228. 23, 25, 26, 76

W. H. Richardson (1972). 'Bayesian-based iterative method of image restoration'. *J. Opt. Soc. Am. A* **62**(1):55–59. 29

P. Rodríguez & B. Wohlberg (2009). 'Efficient minimization method for a generalized total variation functional'. *IEEE Trans. Image Process.* **18**(2):322–332. 36

L. I. Rudin, et al. (1992). 'Nonlinear total variation based noise removal algorithms'. *Phys. D.* **60**:259–268. 34, 35, 36, 129, 152

A. Santos & I. T. Young (2000). 'Model-based resolution: applying the theory in quantitative microscopy'. *Appl. Opt.* **39**(17):2948–2958. 61

P. Sarder & A. Nehorai (2006). 'Deconvolution methods for 3-D fluorescence microscopy images'. *IEEE Signal Process. Mag.* **23**(3):32–45. 22

J. Schlecht, et al. (2006). 'Statistical inference of biological structure and point spread functions in 3D microscopy'. In *Proc. Third International Symposium on 3D Data Processing, Visualization, and Transmission*, pp. 373–380, Chapel Hill, USA. 61

J. W. Shaevitz & D. A. Fletcher (2007). 'Enhanced three-dimensional deconvolution microscopy using a measured depth-varying point-spread function'. *J. Opt. Soc. Am. A* **24**(9):2622–2627. 47, 63, 76

S. L. Shaw (2006). 'Imaging the live plant cell'. *The Plant Journal* **45**(4):573–598. 76, 107, 133

L. A. Shepp & Y. Vardi (1982). 'Maximum likelihood reconstruction for emission tomography'. *IEEE Trans. Med. Imaging* **MI-1**(2):113–122. 59

C. J. R. Sheppard (1986a). 'The spatial frequency cut-off in three-dimensional imaging'. *Optik* **72**(4):131–133. 14

C. J. R. Sheppard (1986b). 'The spatial frequency cut-off in three-dimensional imaging II'. *Optik* **74**(3):128–129. 14

C. J. R. Sheppard (1988). 'Depth of field in optical microscopy'. *J. Microsc.* **149**:73–75. 6

L. Sherman, et al. (2002). 'Adaptive correction of depth-induced aberrations in multiphoton scanning microscopy using a deformable mirror'. *J. Microsc.* **206**(1):65–71. 110

J.-B. Sibarita (2005). 'Deconvolution microscopy'. *Adv. Biochem. Eng. Biotechnol.* **95**:201–243. 29

B. Simon, et al. (2008). 'Tomographic diffractive microscopy of transparent samples'. *Eur. Phys. J. Appl. Phys.* **44**:29–35. 138

J.-L. Starck & A. Bijaoui (1994). ‘Filtering and deconvolution by the wavelet transform’. *Signal Process.* **35**(3):195–211. 37

P. A. Stokseth (1969). ‘Properties of a defocused optical system’. *J. Opt. Soc. Am. A* **59**:1314–1321. 56, 76, 95

A. N. Tikhonov & V. A. Arsenin (1977). *Solution of Ill-posed Problems*. Winston and Sons. 23, 25, 34, 36

T. Tommasi, et al. (1993). ‘3-D reconstruction in optical microscopy by a frequency-domain approach’. *Signal Process.* **32**(3):357–366. 23, 25, 26

P. Török, et al. (1997). ‘The role of specimen-induced spherical aberration in confocal microscopy’. *J. Microsc.* **188**(2):158–172. 134, 138

P. Török, et al. (1995). ‘Electromagnetic diffraction of light focused through a planar interface between materials of mismatched refractive indices: an integral representation’. *J. Opt. Soc. Am. A* **12**(2):325–332. 63

G. M. P. van Kempen & L. J. van Vliet (2000). ‘Background estimation in nonlinear image restoration’. *J. Opt. Soc. Am. A* **17**(3):425–433. 12

G. M. P. van Kempen, et al. (1997). ‘A quantitative comparison of image restoration methods for confocal microscopy’. *J. Microsc.* **12**:354–365. 23, 30

P. J. Verveer, et al. (1999). ‘A comparison of image restoration approaches applied to three-dimensional confocal and wide-field fluorescence microscopy’. *J. Microsc.* **193**:50–61. 23, 32, 34, 36

P. J. Verveer & T. M. Jovin (1997). ‘Efficient superresolution restoration algorithms using maximum a posteriori estimations with application to fluorescence microscopy’. *J. Opt. Soc. Am. A* **14**(8):1696–1706. 34

G. Vicidomini, et al. (2009). ‘Automatic deconvolution of 4Pi-microscopy data with arbitrary phase’. *Opt. Lett.* **34**(22):3583–3585. 44

T. D. Visser, et al. (1992). ‘Refractive index and axial distance measurements in 3-D microscopy’. *Optik* **90**:17–19. 122

C. Vonesch & M. Unser (2007). 'Fast wavelet-regularized image deconvolution'. In *Proc. IEEE International Symposium on Biomedical Imaging*, pp. 608–611, Washington, D.C., USA. 37

Z. Wang & A. C. Bovik (2009). 'Mean squared error: love it or leave it? A new look at signal fidelity measures'. *IEEE Signal Process. Mag.* **26**(1):98–117. 17

P. Weiss (2008). *Algorithmes rapides d'optimisation convexe. Application à la reconstruction d'images et à la détection de changements*. Ph.D. thesis, Université de Nice-Sophia Antipolis. 36

Y. Weiss & W. T. Freeman (2007). 'What makes a good model of natural images?'. In *Proc. IEEE Comp. Vis. and Pat. Recog.*, pp. 1–8, Minneapolis, USA. 31

R. M. Willett, et al. (2003). 'Wavelet-based superresolution in astronomy'. In F. Ochsenbein, M. G. Allen, & D. Egret (eds.), *Proc. Astronomical Data Analysis Software and Systems*, vol. 314 of *Astronomical Society of the Pacific*, pp. 107–116, Strasbourg, France. 37

B. Zhang (2007). *Contributions to fluorescence microscopy in biological imaging: PSF modeling, image restoration, and super-resolution detection.* Ph.D. thesis, l'École Nationale Supérieure des Télécommunications, Paris. 4, 5, 7, 16, 30, 54

B. Zhang, et al. (2007). 'Gaussian approximations of fluorescence microscope point-spread function models'. *Appl. Opt.* **46**(10):1819–1829. 10, 61, 62, 84, 85, 101