Parametric Blind Deconvolution for Confocal Laser Scanning Microscopy (CLSM)-Proof of Concept

Praveen Pankajakshan, Laure Blanc-Féraud, B. Zhang, Zvi Kam,
Jean-Christophe Olivo-Marin, Josiane Zerubia

To cite this version:

Praveen Pankajakshan, Laure Blanc-Féraud, B. Zhang, Zvi Kam, Jean-Christophe Olivo-Marin, et al.. Parametric Blind Deconvolution for Confocal Laser Scanning Microscopy (CLSM)-Proof of Concept. [Research Report] RR-6493, 2008. inria-00269265v1

HAL Id: inria-00269265
https://inria.hal.science/inria-00269265v1

Submitted on 2 Apr 2008 (v1), last revised 4 Apr 2008 (v2)

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Parametric Blind Deconvolution for Confocal Laser Scanning Microscopy (CLSM)-Proof of Concept

Praveen PANKAJAKSHAN, Laure BLANC-FÉRAUD, Bo ZHANG, Zvi KAM,

Jean-Christophe OLIVO-MARIN and Josiane ZERUBIA

N° 1234

April 2008

Thème COG

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--1234--FR+ENG

Parametric Blind Deconvolution for Confocal Laser Scanning Microscopy (CLSM)-Proof of Concept

Praveen PANKAJAKSHAN*, Laure BLANC-FÉRAUD*, Bo ZHANG[†], Zvi KAM[‡], Jean-Christophe OLIVO-MARIN[†] and Josiane ZERUBIA*

Thème COG — Systèmes cognitifs
Projet ARIANA

Rapport de recherche n° 1234 — April 2008 — 45 pages

Abstract: We propose a method for the iterative restoration of fluorescence Confocal Laser Scanning Microscope (CLSM) images with parametric estimation of the acquisition system's Point Spread Function (PSF). The CLSM is an optical fluorescence microscope that scans a specimen in 3D and uses a pinhole to reject most of the out-of-focus light. However, the quality of the image suffers from two primary physical limitations. The first is due to the diffraction-limited nature of the optical system and the second is due to the reduced amount of light detected by the photomultiplier tube (PMT). These limitations cause blur and photon counting noise respectively. The images can hence benefit from post-processing restoration methods based on *deconvolution*. An efficient method for parametric blind image deconvolution involves the simultaneous estimation of the specimen 3D distribution of fluorescent sources and the microscope PSF. By using a model for the microscope image acquisition physical process, we reduce the number of free parameters describing the PSF and introduce constraints. The parameters of the PSF may vary during the course of experimentation, and so they have to be estimated directly from the observation data. We also introduce a priori knowledge of the specimen that permits stabilization of the estimation process and favorizes the convergence. Experiments on simulated data show that the PSF could be estimated with a higher degree of accuracy and those done on real data show very good deconvolution results in comparison to the theoretical microscope PSF model.

Key-words: Confocal Laser Scanning Microscope (CLSM), Bayesian restoration, blind deconvolution, Point Spread Function (PSF), Richardson-Lucy (RL), Total Variation (TV), parameter estimation

---

* ARIANA Project-team, CNRS/INRIA/UNSA, 2004 Route des lucioles-BP93, 06902 Sophia-Antipolis Cedex, France

[†] Quantitative Image Analysis Unit, Institut Pasteur, 25-28 rue du Docteur Roux, 75015 Paris, France

[‡] Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel 76100.

Déconvolution aveugle paramétrique en imagerie de Microscopie Confocale à Balayage Laser (CLSM).

Résumé : Nous proposons une méthode de restauration itérative d'images de fluorescence CLSM et d'estimation paramétrique de la fonction de flou (PSF) du système d'acquisition. Le CLSM est un microscope qui balaye un échantillon en 3D et utilise une sténopée pour rejeter la lumière en dehors du point de focalisation. Néanmoins, la qualité des images souffre de deux limitations physiques. La première est due à la diffraction due au système optique et la seconde est due à la quantité réduite de lumière détectée par le tube photo-multiplicateur (PMT). Ces limitations induisent respectivement un flou et du bruit de comptage de photons. Les images peuvent alors bénéficier d'un post-traitement de restauration fondé sur la déconvolution. Le problème à traiter est l'estimation simultanée de la distribution 3D de l'échantillon des sources fluorescentes et de la PSF du microscope (i.e. de déconvolution aveugle). En utilisant un modèle de processus physique d'acquisition d'images microscopiques (CLSM), on réduit le nombre de paramètres libres décrivant la PSF et on introduit des contraintes. On introduit aussi des connaissances a priori sur l'échantillon ce qui permet de stabiliser le processus d'estimation et de favoriser la convergence. Des expériences sur des données synthétiques montrent que la PSF peut être estimée avec précision. Des expériences sur des données réelles montrent de bons resultats de déconvolution en comparaison avec le modèle théorique de la PSF du microscope.

Mots-clés : Microscopie Confocale à Balayage Laser (CLSM), restauration Bayésienne, déconvolution aveugle, la fonction de flou (PSF), Richardson-Lucy (RL), Variation Totale (TV), estimation paramétrique

Acknowlegement

This research was funded by the P2R Franco-Israeli Collaborative Research Program[1]. The authors gratefully acknowledge Dr. Caroline Chaux (Université Paris-Est, France), Mr. Gilbert Engler (INRA Sophia-Antipolis, France) and Prof. Arie Feuer (Technion, Israel) for several interesting discussions. We would also like to thank INRIA for supporting the Ph.D. of the first author through a CORDI fellowship and CNRS for supporting the Ph.D. of the third author. Additionally, our sincere gratitude goes to Prof. Stéphane Noselli and Mrs. Fanny Serman (Ph.D. candidate) from the Institute of Signaling, Development Biology & Cancer UMR 6543/CNRS/UNSA for the images presented in Fig. 12 of Section 4.2.3, and the useful comments for the validation.

---

[1] See http://www-sop.inria.fr/ariana/Projets/P2R/index.html for more information

Contents

Abbreviations

3D	Three-Dimensional
AM	Alternate-Minimization
AU	Airy Units
CCD	Charge-Coupled Device
CG	Conjugate-Gradient
CLSM	Confocal Laser Scanning Microscope
EM	Expectation Maximization
ENR	Edge-to-Noise ratio
FWHM	Full-Width at Half Maximum
GEM	Generalized Expectation Maximization algorithm
GFP	Green Fluorescent Protein
iff	if and only if
JVC	Janson Van Cittert
LLS	Linear Least-Squares
LSM	Laser Scanning Microscope
MAP	Maximum a Posteriori
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimate
MRF	Markov Random Fields
MSE	Mean Squared Error
NA	Numerical Aperture
NLS	Nonlinear Least-Squares
OTF	Optical Transfer Function
PBD	Parametric Blind Deconvolution
pdf	Probability Density Function
PMT	Photomultiplier Tube
PSF	Point Spread Function
RL	Richardson-Lucy algorithm
SNR	Signal-to-Noise Ratio
TM	Tikhonov-Miller
TV	Total Variation
WFM	Wide-Field Microscope
w.r.t	with respect to

Notations

$\hat{(\cdot)}$	Estimate
c	Clique in the set of cliques C
$\delta(\cdot)$	Dirac-delta function
ϵ	Convergence criterion
$\mathcal{F}(\cdot)$	Fourier Transform
h	Point-Spread Function (PSF) of the microscope
i	Observed image
k	Iteration index of the parameter estimation algorithm
λ	Regularization parameter or Average Photon Flux (as the case maybe)
λ_{em}	Emission wavelength
λ_{ex}	Excitation wavelength
$\mathcal{L}(\cdot)$	Cost function
μ_{med}	Refractive index of a medium
(μ, σ^2)	Mean and Variance of a normal distribution $\mathcal{N}$
$n(\mathbf{x})$	Additive gaussian noise
n	Iteration index of the deconvolution algorithm
o	Specimen or object imaged
$\mathcal{P}(\cdot)$	Poisson distribution
$P(\cdot\|\cdot)$	Conditional Probability
P	Pupil function
(r, ϕ, z)	Cylindrical coordinates
θ	Parameter vector to be estimated
$Tr(A)$	Trace of a matrix A
$\mathbf{u}$	Vector of coordinates in the frequency space
$\mathbf{x}$	Vector of coordinates in image space (2D or 3D)
z	Axial distance in the image space

1 Introduction

There are many problems in science and engineering that involve determination of the internal structure of a system from just a few external measurements. Unlike the direct approach, these are *ill-posed* in the Hadamard sense [1] that the solution either does not exist, is not unique or is unstable with perturbations in the input data. Image restoration that aims at reversing a system's degradation also suffers from a similar restriction. Physical imaging systems generally fail in transmitting all the information about an object, and worse sometimes adding artifacts to the output image. The task of restoration is difficult because in microscopy, as with many other imaging techniques, different possible configurations of the imaging object is possible given only the observation and little or no information of the imaging system.

1.1 Organization of this report

Before we introduce the reader to our actual problem of interest, we devote a section on Fluorescence Microscopes that looks back at the differences between a conventional Wide-Field Microscope (WFM) and the Confocal Laser Scanning Microscope (CLSM). Although the scope of this report is restricted to restoration of images acquired by the CLSM, the WFM is also familiarized to compare both their imaging capabilities. We then discuss their individual response characteristics to a point source of light that is imaged under specific conditions. Due to the limitation of the photon counting proces, the output image obtained from the microscope is often corrupted by noise. The nature of the noise, its mathematical modeling and handling, prior work on this subject, and deconvolution is acquainted in Section 2. Section 3 in its entirety is dedicated to our proposal on the joint restoration and estimation of the image of the object and the microscope response function using a Bayesian methodology. Direct restoration from the observed data is very difficult, and hence it is necessary to define the respective underlying models for the object and the response function. An Alternate Minimization (AM) algorithm is then recommended to solve this particular problem. This AM algorithm was then tested on images of phantom objects and real data, the results of which are presented in Section 4. We then conclude in Section 5 with comments on this work and a discussion on the future work.

1.2 Fluorescence Microscope

Fluorescence is a phenomenon where electrons, when excited by light of a certain wavelength, emit light of a longer wavelength. This is because the relaxation of the electron from the higher to a lower state causes emission of photon that has lesser energy than the photon that was absorbed. This energy difference, $\Delta E \propto 1/\lambda_{em}$, is known as the *Stokes shift*. The wavelength of the emitted light is thus always *red-shifted* (i.e. is nearer to the red light in the spectrum), in comparison to the excitation light wavelength. The greater the Stokes shift, the easier it is to separate the excitation light from the emitted light. The selection of excitation and emission wavelengths are controlled by appropriate filters. The emission curve is often a mirror image of the excitation curve. Fig. 1 shows the emission

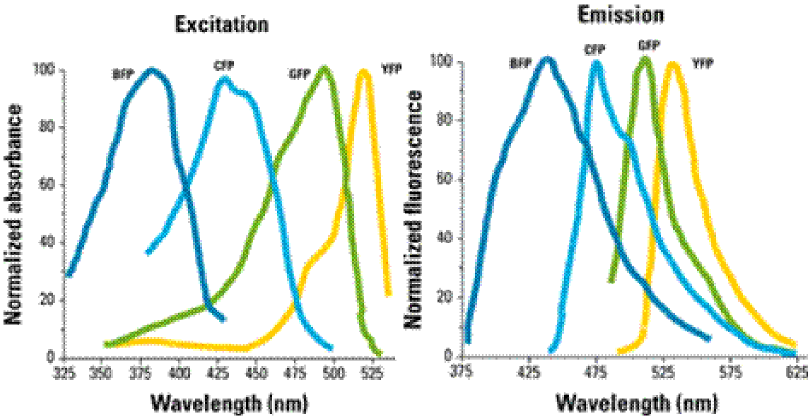


Figure 1: Excitation and emission spectra for the various fluorescence proteins

and excitation spectra for some common fluorescent labeling proteins.

In fluorescence microscopy, a high intensity monochromatic laser light is used for illuminating a sample. Each point in the sample acts as a secondary light source either naturally or because of the expression of the fluorescent labeling proteins. The molecule that has the ability to fluoresce is called a "*fluorophore*" (atomic compound responsible for fluorescence) or fluorochrome (usually a dye that renders a body fluorescent). To achieve maximum fluorescence intensity, the fluorochrome is usually excited at the wavelength at the peak of the excitation curve, and the emission is selected at the peak wavelength (or other wavelengths chosen by the observer) of the emission curve. For example, when excited with a monochromatic source of 488nm, the "*Green Fluorescent Protein*" (GFP) has an emission wavelength with a peak at 520 nm (and quantum efficiency of 0.8).

1.2.1 Incoherent Point Spread Function (PSF)

The optics of an observation system allows inspection of a specimen, but the image finally obtained is often not perfect. This is an intrinsic limit of any imaging system and is the determining factor in assessing the system's resolution limit. This can be justified by diffraction, and an image whose resolution is thus limited is considered to be *diffraction-limited* [2] [3].

Mathematically the process of degradation is characterized by the impulse response (by imaging point-like source of light) or PSF, which models the propagation and recording of the electromagnetic radiation from a point source. The PSF thus displays a radial diffractive ring pattern (expanding with defocus) that is introduced by the finite-lens aperture. Each optical section has the in-focus plane and also out-of-focus contributions from other parts of the object. Due to the random nature of the photon emission, fluorescence microscopy is an incoherent imaging process. That means, each point of the sample contributes independently to the light intensity distribution in the image space.

Optical theory includes sophisticated models of this blurring and with modern computational capacity, we can apply such models to restoration. The nature of the PSF for fluorescence microscope has been researched extensively both theoretically and experimentally (see Section 2.4). Because of its importance in deconvolution, a theoretical model of blur or the Point Spread Function (PSF) is introduced in greater detail in Section 3.1. The advantage of such a model is its efficient representation with a few parameters.

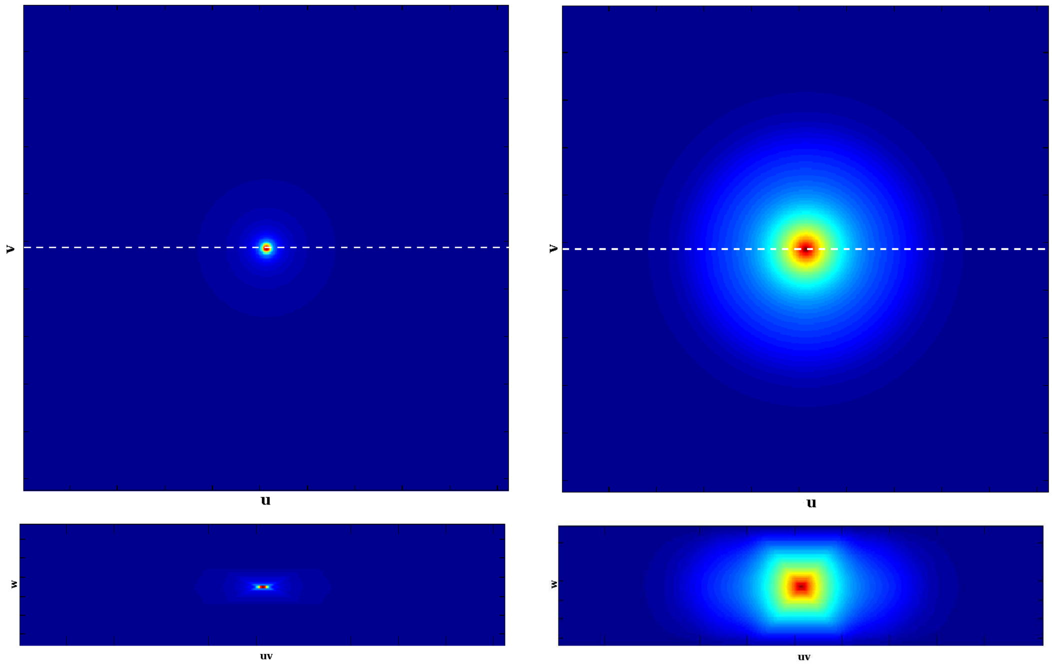
1.2.2 Wide-Field Microscope (WFM)

In the conventional WFM, the entire specimen is illuminated uniformly with an extended light source. When specimens are imaged using such a microscope, each optical section obtained by this focus series has the in-focus plane plus out-of-focus contributions from other parts of the object that obscure the image and causes severe axial blur. This is because secondary fluorescence emitted by the sections that are away from the region of interest often interferes with the resolution of those features that are in focus. This situation is especially serious for specimens having a thickness greater than about $2\mu m$. Additionally there is a severe degradation to the image caused by the diffraction at the objective lens.

The object's maximum intensity value decreases only as z^{-2}, with z being the distance in axial direction from the focus. Even though there are several advantages of WFM over CLSM, like faster acquisition time or increased signal strength, the off-focus suppression is insufficient and they suffer from something called the "*missing cone of frequencies*" problem along the optical axis [4]. Thus, objects with many frequencies in that region will be poorly imaged. For example, a thin fluorescent sheet, oriented along the focal plane has its frequencies in that region. Thus, this homogeneous horizontal plane can not be imaged or located. Since the total (plane-wise integrated) intensity remains constant objects below or above the layer always contribute to the final image of that layer. In Fig. 2, we illustrate this problem for a microscope setting with an excitation wavelength peak $\lambda_{em} = 488$ nm, emission wavelength of 520nm, and pixel sizes of $94 \times 94 \times 283$nm for the WFM and $43 \times 43 \times 130$nm for the CLSM. Fig. 6(b) shows the radial and axial slices of a wide-field microscope PSF generated using the *Huygens*© software and the respective microscope parameters.

1.2.3 Confocal Laser Scanning Microscope (CLSM)

The CLSM (also referred to sometimes as the Laser Scanning Confocal Microscope (LSCM)) is an optical fluorescence microscope that scans a biological specimen in 3D using a focused laser spot and uses a pinhole before the detection to reject out-of-focus fluorescence [5]. Fig. 3 shows the schematic of a typical confocal laser scanning microscope used for scanning biological specimens in 3D. CLSM focuses the objective at different depths inside a



(a) (b)

Figure 2: Illustration of the missing-cone problem in the OTF of a (a) WFM and its absence in a (b) CLSM (top: axial plane, bottom: radial plane)

specimen allowing the 3D visualization of cells, tissues and embryos without a need for physical sectioning. The CLSM does not suffer from the *"missing cone frequency"* problem as in the WFM case and the sensitivity decreases approximately as z^{-4}. The difference between the application of the CLSM and WFM is that the former is suited for thick specimens such as embryos and tissues while the latter is a powerful method for imaging samples requiring low light levels such as living cells and nucleic acids.

Although the out-of-focus contribution in confocal images is greatly reduced, it is not totally eliminated, and in practice is highly dependent on the pinhole diameter. This reduces the maximal resolution obtainable by the imaging system and contrast, and often complicates the quantitative analysis of the 3D specimen.

In general the observation can be modeled as:

$$i = \varphi(B(o)) \odot n \tag{1}$$

where o is the object or specimen under inspection, i is the observed image, B models the blurring effect produced, n is the noise, $\odot$ is an operator of addition or multiplication and φ is a transformation which can be linear, non linear, deterministic or probabilistic.

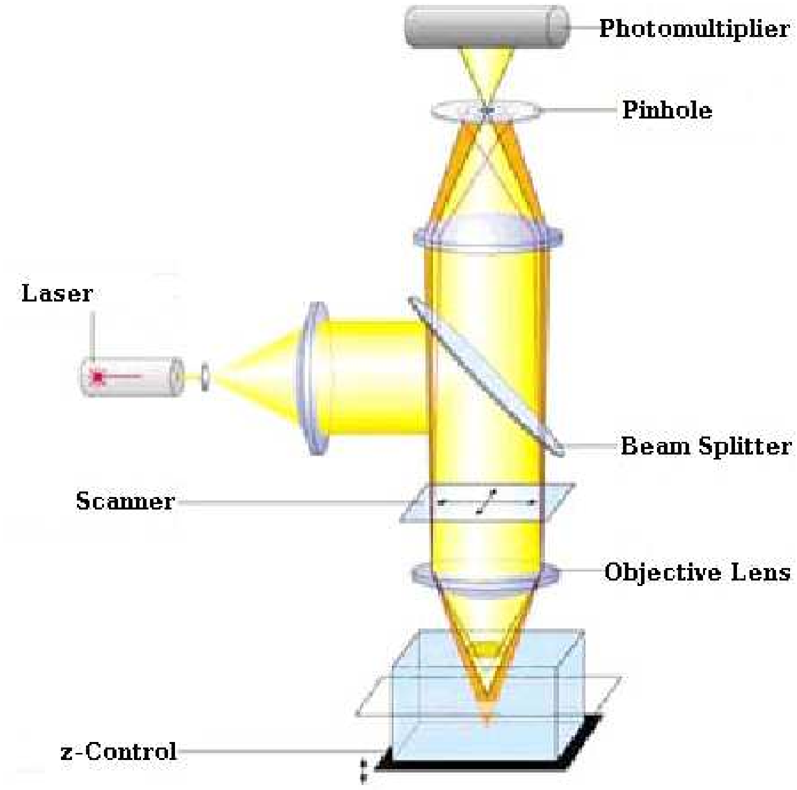


Figure 3: Schematic of a CLSM showing the in-focus and out-of-focus planes

1.3 Problem Formulation

The main difficulty in restoring the 3D image is that the exact PSF is not precisely known, denoising the image can induce artifacts and restoration by *deconvolution* is an ill-posed problem.

Computationally, deconvolution uses the PSF to remove the out-of-focus light that was described earlier, and thus improves the contrast and the resolution. Restoration can be done using either *blind* or *non-blind* techniques. The difference between them is that in the first case the PSF is known a priori, while in the second case the PSF is estimated along with the data itself. There are three kinds of methods to obtain the microscope PSF for the deconvolution process. Firstly, in the experimental approach [4], point-like objects in the specimen are imaged and processed to obtain the PSF. The experimental PSF in itself might have low contrast (therefore may be recorded only at a finite range of defocus values) or might be contaminated by noise, and it requires point-like sources in every image. In addition, PSF measured in one sample (typically fluorescent micro-beads stuck to a cover slide) may not represent the exact PSF applicable for another sample (such as live cells in physiological buffers). Secondly, we can use an analytical model of the PSF [6] with the acquisition system's physical information as parameters. However, in this case, the optical parameters of the set up (for example small residual phase aberrations in the objective) are not known or might change during the course of an experimentation (for example, due to heating of live samples). In the third case, the blind deconvolution approach that we propose, estimates the specimen and the unknown PSF parameters from the observed image simultaneously. It finds the closest approximation of the specimen reconstructed structure without requiring experimental or analytical calculation of the optical system PSF.

The scope of this report is restricted to restoring images and estimating the parameters of a diffraction-limited PSF with no aberrations and is spatially-invariant. The advantage of a spatially-invariant PSF is that the degradation model is a convolution. It is important to note that the spatial-invariance assumption is not true when there are aberrations in the microscope. These aberrations may come from a variety of different sources: refractive index gradients in the specimen material, refractive index mismatch, bad alignment and others. H. E. Keller [2] provides a survey of all known types of PSF aberrations. In practice, the spherical aberration [7] is the most predominant among them all. Here there is an axial asymmetry in the shape of the PSF, with a corresponding increase in spread particularly along the optical axis. The most common cause of spherical aberration is a mismatch between the refractive indices of the lens immersion medium and the mounting medium in which the specimen rests.

2 Review of prior work

The review of the prior work on the subject of restoration led us to conclude that essentially most of them differ primarily in the models they choose to described the observed phenomenon. As a consequence, the approach that each of them adopt for the restoration becomes different. When both the observation model and the approach are chosen to be the same, they differ by their assumption of PSF models. We will try to explain more clearly and highlight these differences and point out the advantages and drawbacks of each method. This will prepare the necessary background in understanding the proposed approach in the next section.

2.1 Noise Sources and Models

Noise is a quasi-random disarrangement of detail in the observed image. It is quasi-random in nature because the statistical distribution of noise can be predicted if the mechanics of its source are known.

2.1.1 Sources of noise

The principal instruments used for fluorescence detection are CCDs and PMTs. In digital microscopy, the source is either the signal itself (so-called photon shot noise), or the digital imaging system. The mechanics of both sources are understood and the statistical distribution of noise is therefore known. Signal dependent noise is characterized by a Poisson distribution while imaging system noise usually follows a Gaussian distribution. The interested reader may refer to [8] and [9] for a more detailed discussion on the various types of noises present in fluorescence microscopes. Here we will introduce briefly the noises that seem significant for our current work.

Photon noise Supposing the average photon flux to be λ_p, the observed photon number N_p follows a Poisson distribution with intensity parameter λ_p, i.e., $N_p \sim P(\lambda_p)$. If we dispose of a high photon flux, N_p will be asymptotically normally distributed with both the mean and the variance equal to λ_p.

Dark Noise The kinetic vibration of silicon atoms in the CCD substrate will liberate electrons or holes even when no incident fluorescence photon is present. The resulting charge will contribute to the final signal and is termed dark noise. Secondary sources of the dark noise are cosmic rays or external high-energy radiation from nearby sources such as an indoor illumination. The dark noise N_d follows also a Poisson distribution, i.e., $N_d \sim P(\lambda_d)$, where λ_d represents the average dark flux. In practice, the CCDs are usually cooled to reduce the dark noise.

2.1.2 Noise Models

Gaussian assumption Classical restoration approaches differ in the initial models they choose for the observed phenomenon. Most often they assume that an additive gaussian noise observation model is sufficient to describe the noise characteristics of the fluorescent microscope [10] [11].
If $o(\mathbf{x})$ and $i(\mathbf{x})$ (assumed to be bounded and positive) denote the original and observed images respectively, the model in (1) for the Gaussian assumption can be expressed as,

$$i(\mathbf{x}) = [h * o](\mathbf{x}) + n(\mathbf{x}),\ \mathbf{x} \in \Omega \tag{2}$$

where $n(\mathbf{x}) \sim \mathcal{N}(0, \sigma_n^2)$ is the additive Gaussian white noise, $h(\mathbf{x})$ is the microscope PSF. The object's intensity is given by $\{o(\mathbf{x}) : \mathbf{x} \in \Omega\}$. Where for the discrete image, the domain $\Omega = (x, y, z) : 0 \leq x \leq N_x - 1, 0 \leq y \leq N_y - 1, 0 \leq z \leq N_z - 1$ is a finite lattice and defines the region over which the intensity is non-vanishing and is finite in nature. If we assume that the imaging system is linear and shift-invariant, $*$ denotes the convolution in 2D or 3D as the case may be. This observation is often not valid especially in the case of photon counting process where the noise characteristics is close to a Poissonian process.

Poisson assumption The detector of the fluorescence microscope behaves ideally as a photon counter (ignoring electronic amplification noise). For low illumination conditions, the number of photons reaching the detector is small, and the statistical variation in the number of detected photons can be described by a Poisson process [12] (see [9] for more details). It is however important to note that the Poissonian model is valid only for low-intensity regions where the photomultiplier functions in the counting state.
For a Photomultiplier Tube (PMT) operating in the photon-counting mode, the detected signal S will be (no readout noise),

$$S = N_p + N_d \sim P(\lambda_p + \lambda_d) \tag{3}$$

where, N_p, N_d, λ_p and λ_d are as defined earlier.
The image degradation model (see Fig. 4) can be expressed as:

$$i(\mathbf{x}) = \mathcal{P}([h * o](\mathbf{x})) \tag{4}$$

where $*$ denotes 3D convolution (assuming linearity property of the imaging system), $\mathcal{P}(\cdot)$ denotes a Poissonian process, $h(\mathbf{x})$ is defined in (35), and

$$(h * o)(\mathbf{x}) = \sum_{\mathbf{x}' \in \Omega} h(\mathbf{x} - \mathbf{x}')o(\mathbf{x}') \tag{5}$$

The likelihood of X is modeled as [12]:

$$P(Y = i|X = o) = \prod_{\mathbf{x} \in \Omega} \frac{[h * o](\mathbf{x})^{i(\mathbf{x})} e^{-[h*o](\mathbf{x})}}{i(\mathbf{x})!} \tag{6}$$

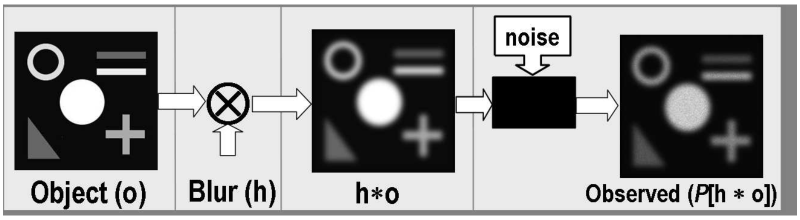


Figure 4: Schematic model of the degradation process.

where the mean of the Poisson process is given by $[h * o](\mathbf{x})$. In [13] a background term was also introduced in the above model, and the experimental results highlighted the lack of improvement on the restoration results by inclusion of the background signal. Poisson noise, unlike Gaussian noise is signal intensity dependent, which makes separation of the signal from the noise a very difficult task.

Mixed Noise In the case of Widefield Microscopes, often the CCD has a dominating read-out noise, the signal model then becomes,

$$\begin{aligned} S &= \alpha(N_p + N_d) + N_r \\ S &\approx \alpha(\lambda_p + \lambda_d) + B,\, \sigma \gg \alpha \end{aligned} \tag{7}$$

where α is the overall camera gain, N_r is the read-out noise and $B \sim N(\mu, \sigma^2)$

2.2 Deconvolution algorithms

There are many denoising techniques for removing the noise (see [9] [14] [15]) that is present in the biological images. However, it is difficult to analytically model the resulting denoised image as a function of the original object and the PSF, and then retrieve the object by deconvolution. Instead of relying on sequential algorithms to remove the noise and the blur, we use a simultaneous approach based on regularization models (see Section 2.3).

The deconvolution algorithms available in literature are summarized in Table 1. In general, they minimize an energy function of the form (with the exception of [16] and [17] where no observation model is assumed):

$$\mathcal{L}(o) = \mathcal{L}_{obs}(o) + \mathcal{L}_{reg}(o) \tag{8}$$

Here, $\mathcal{L}_{obs}(o)$ is a measure of fidelity to the data and $\mathcal{L}_{reg}(o)$ is the penalty term on the object estimation.

Accordingly, we have divided the following few sections based on the method employed for solving the observation model described earlier in Section 2.1.2 and the object model. A simplest approach to deblurring is using the Nearest-Neighbors [17] which assumes that

METHOD	REFERENCES
No-neighbors	[16]
Nearest Neighbors	[17]
Linear Methods	
Weiner, inverse filter	[18] [19]
Linear Least squares (LLS)	[20]
Tikhonov	[21]
Constrained Iterative	
Jansson Van Cittert	[17]
Nonlinear least squares	[22]
Statistical	
Maximum Likelihood	[23]
Maximum A Posteriori	[24] [25] [26] [27] [28]

Table 1: Deconvolution Algorithms

the most blurring within a slice is the light scattered from its two neighboring slices. The intra-plane PSF is normally much smaller than the inter-plane PSF. Using these assumptions, and if $\hat{O}(\mathbf{u}, z)$ and $I(\mathbf{u}, z)$ are the 2D Fourier transforms of the estimated specimen and the acquired image at the z^{th} slice,

$$\hat{O}(\mathbf{u}, z) = I(\mathbf{u}, z) - \lambda[\hat{H}_1(\mathbf{u})I(\mathbf{u}, z-1) + \hat{H}_1(\mathbf{u})I(\mathbf{u}, z+1)] \tag{9}$$

where $\hat{H}_1(\mathbf{u})$ and $\hat{H}_1(\mathbf{u})$ are the inter-plane PSF for two neighboring planes respectively. λ adjusts the contribution of the two neighboring planes to the central plane.

2.2.1 Linear Methods

Inverse Filtering In the absence of a noise model, the image distortion by the PSF could be reversed by using an inverse filter $h^{\dagger}(\mathbf{x})$ such that,

$$\hat{o}(\mathbf{x}) = \mathcal{F}^{-1}(H^{\dagger}(\mathbf{u})I(\mathbf{u})) \tag{10}$$

where $H^{\dagger}(\mathbf{u})$ and $I(\mathbf{u})$ are the 3-D Fourier Transforms of the inverse filter and the observation respectively. $\mathcal{F}^{-1}$ is the inverse Fourier transform operation. Directly inverting the PSF leads to an inverse filter $H^{\dagger}(\mathbf{u}) = H^{-1}(\mathbf{u})$. Under practical situations however, at high spatial frequencies, the inversion suffers from noise amplification. This problem can be tackled by using a truncated inverse filter wherein the higher frequencies are bounded by a small positive constant ϵ as:

$$\hat{o}(\mathbf{x}) = \begin{cases} \mathcal{F}^{-1}(H^{\dagger}(\mathbf{u})I(\mathbf{u})), & \text{if } |H(\mathbf{u})| \geq \epsilon \\ 0, & \text{otherwise} \end{cases}$$

Linear least squares filtering Similar to the inverse filter, if we assume no model on the noise, then the estimate of the object is obtained as:

$$\hat{o}(\mathbf{x}) = (H^{*}R^{-1}H)^{-1}H^{*}R^{-1}i(\mathbf{x}) \tag{11}$$

and the MSE is given as:

$$\mathbb{E}(||o(\mathbf{x}) - \hat{o}(\mathbf{x})||^{2}) = Tr[(H^{*}R^{-1}H)^{-1}] \tag{12}$$

H^{*} is the complex conjugate of the PSF H, R is the covariance matrix of noise n with $\mathbb{E}(n) = 0$. If n is white uncorrelated noise, $R = \sigma^{2}I$ and

$$\hat{o}(\mathbf{x}) = (H^{*}H)^{-1}H^{*}i(\mathbf{x}) \tag{13}$$

and the MSE is given as:

$$\mathbb{E}(||o(\mathbf{x}) - \hat{o}(\mathbf{x})||^{2}) = \sigma^{2}Tr[(H^{*}H)^{-1}] \tag{14}$$

(13) is similar to (10) with $H^{\dagger} = (H^{*}H)^{-1}H^{*}$.
In order to improve the resolution and the contrast, Projection onto convex sets (POCS) was proposed as an alternative. POCS is a powerful mathematical tool proposed by Bregman [29] and later introduced to deconvolve microscopic images in [30].

Wiener Filtering If the signal-independent additive Gaussian noise is considered, for a known $h(\mathbf{x})$, $i(\mathbf{x})$ is deconvolved as:

$$o(\mathbf{x}) = i(\mathbf{x}) * h^{\dagger}(\mathbf{x}) \tag{15}$$

In the frequency domain, $h^{\dagger}(\mathbf{x})$ is given as:

$$H^{\dagger}(\mathbf{u}) = \frac{H^{*}(\mathbf{u})}{(|H(\mathbf{u})|^{2} + (\frac{P_{n}(\mathbf{u})}{P_{o}(\mathbf{u})}))} \tag{16}$$

where $H(\mathbf{u})$ is the Fourier transform of $h(\mathbf{x})$, $P_{n}(\mathbf{u})$ and $P_{o}(\mathbf{u})$ are the power spectral densities of the noise and the object respectively.

2.2.2 Nonlinear Methods

Constrained Iterative method In the Janson Van Cittert (JVC) algorithm, at each 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.

$$\hat{o}^{k+1}(\mathbf{x}) = \hat{o}^{k}(\mathbf{x}) + r\{\hat{o}^{k}(\mathbf{x})\}[i(\mathbf{x}) - (h * \hat{o}^{k})(\mathbf{x})],\ k \geq 1 \tag{17}$$

where, $r\{\hat{o}^k(\mathbf{x})\} = r_0[1 - \frac{2}{b}|\hat{o}^k(\mathbf{x}) - \frac{b}{2}|]\}$, r_0 and b are constants, and

$$\hat{o}^1(\mathbf{x}) = i(\mathbf{x}) + r\{i(\mathbf{x})\}[i(\mathbf{x}) - (h * i)(\mathbf{x})] \tag{18}$$

This method amplifies 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.
D. A. Agard's modification to the algorithm for faster convergence is as follows:

$$\hat{o}^{k+1}(\mathbf{x}) = \hat{o}^k(\mathbf{x}) + r\{\hat{o}^k(\mathbf{x})\}[(i * h)(\mathbf{x}) - (h * h * \hat{o}^k)(\mathbf{x})],\ k \geq 1, \tag{19}$$

and

$$\hat{o}^1(\mathbf{x}) = (i * h)(\mathbf{x}) + r\{(i * h)(\mathbf{x})\}[i(\mathbf{x}) - (h * i * h)(\mathbf{x})] \tag{20}$$

Carrington et al. [22], 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. In spite of these modifications, it was observed that the convergence of these algorithms is dependent on the parameters r_0 and b chosen.

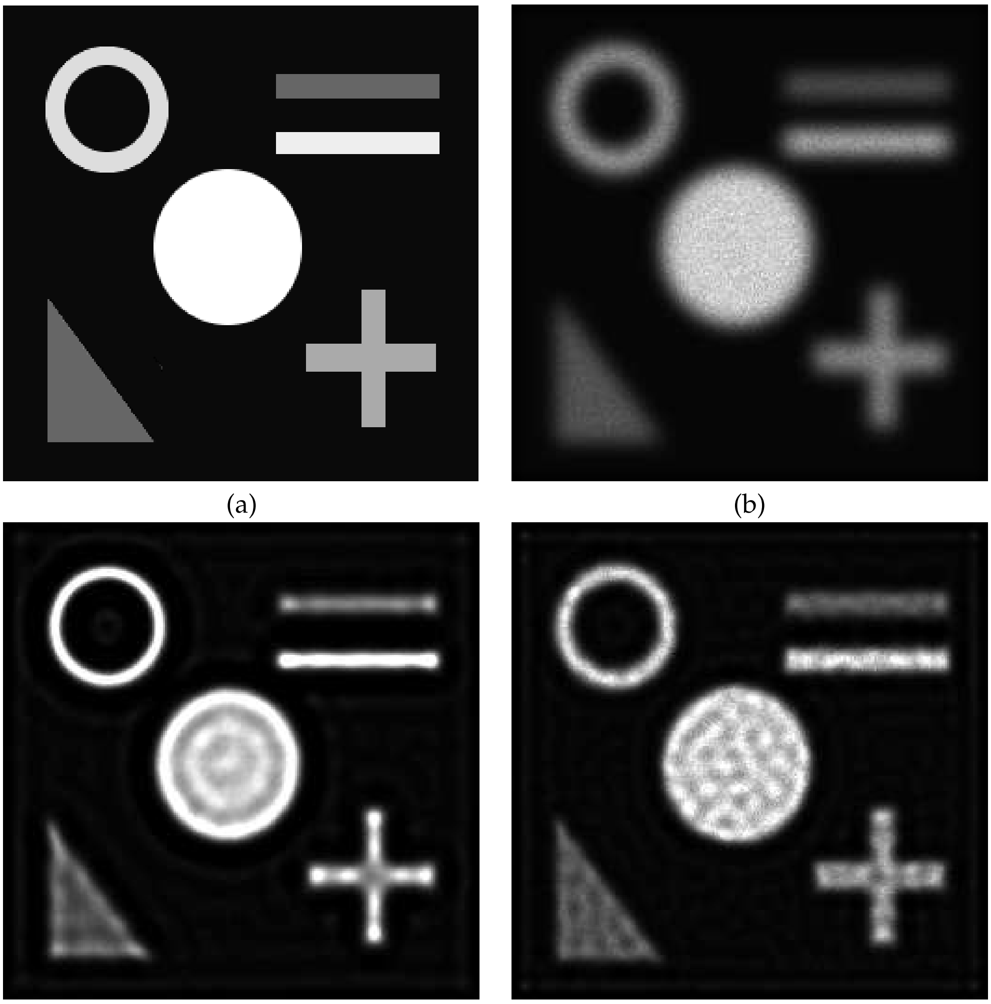
2.2.3 Statistical Methods

Statistical methods are extremely effective when the noise in the acquired 3-D image is fairly strong. These have a more subtle noise handling strategy than the simple regularization. They are also helpful in obtaining certain information not captured by the microscope optics. However, they are more complex and computationally more intensive than the linear and nonlinear methods.

Maximum-Likelihood (ML) If we assume that the background noise is absent,then the maximum-likelihood reconstruction as described by [23] can be written as:

$$\mathcal{L}_{obs}(o(\mathbf{x})) = \begin{cases} ||(h * o)(\mathbf{x}) - i(\mathbf{x})||^2, & \text{Gaussian case} \\ \sum\limits_{\mathbf{x} \in \Omega} [h * o](\mathbf{x}) - i(\mathbf{x}) \log[h * o](\mathbf{x}), & \text{Poissonian case} \end{cases} \tag{21}$$

The ML estimation can also be thought of as a specific form of the Maximum a Posteriori (MAP) estimation method (discussed next) when the a priori model is uniform. For the additive Gaussian noise model of (2), the ML solution is essentially the LLS solution for a known $h(\mathbf{x})$ and assuming that the noise is signal independent. There are many ways of arriving at (21) ([31] looks at it from a statistical angle).



(a) (b)

(c) (d)

Figure 5: An example of ML estimate with no smoothness constraints on a (a) phantom object, and (b) as observed by a CLSM. Assuming that the PSF is known, as the iteration increases, the blur is removed but artifacts start appearing. For example, the central slice after (c) 200 and (d) 500 iterations.

Maximum a Posteriori This section on MAP is a precursor to our analysis of this method in Section 3. Since iterative ML methods do not ensure any smoothness constraints, they thus evolve to a solution that displays many artifacts from noise amplification Fig. 5. We have adopted the MAP method for solving our problem with a prior model on the specimen and the PSF. So,

$$P(o|i) \propto P(i|o)P(o) \tag{22}$$

where $P(i|o)$ is the likelihood of observing the object as given in (6) and $P(o)$ is the knowledge of the object to be discussed next.

2.3 A priori object models

Let $\mathcal{I}(\Omega) = \{o = (o_{ijk}) : \Omega \rightarrow \mathbb{R} | o_{ijk} \in \mathbb{R}\}$ denote all possible images on Ω. Then the ensemble model of an image class refers to any probability distribution $P(o)$ on the image space $\mathcal{I}$ given by the following form:

$$P(o) = \frac{1}{Z} e^{-\lambda E[o]}, \tag{23}$$

where $E[o]$ is a generalized energy and its dual parameter is λ.
We associate with each site $(i, j, k) \in \Omega$ of the object a unique neighborhood $\eta_{ijk} \subseteq \Omega \setminus (i, j, k)$ and the collection of all neighbors $\eta = \{\eta_{ijk} | (i, j, k) \in \Omega\}$ is the neighborhood system. If we assume that the random field $(O = o)$ on a domain Ω is Markovian w.r.t the neighborhood system η, then

$$P(o_{ijk} | o_{\Omega \setminus i,j,k}) = P(o_{ijk} | o_{\eta_{ijk}}) \tag{24}$$

o is a Markov Random Field (MRF) on (Ω, η) if o denotes a Gibbs ensemble on Ω and the energy is a superposition of cliquish potentials

$$E(o) = \sum_{C \in \mathcal{C}} V_C[o] \tag{25}$$

2.3.1 ℓ_2 regularization

Gaussian prior The simplest quadratic penalty on the object is the Gaussian density function. We can write $P(o)$, the pdf of o as, $P(o) = Z^{-1} \exp(-\frac{1}{2\sigma^2} ||o||^2)$, where σ^2 is the variance of the prior Gaussian distribution, and Z is a normalizing constant. This in combination with the Gaussian likelihood function (described previously) gives a Tikhonov type of functional [32].

Tikhonov-Miller Prior Tikhonov-Miller [21][33] introduced a regularization term also based on the ℓ_2 norm of the image. For this case, we can write the statistics on the object as $P(o) = Z^{-1} \exp(-\lambda \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|^2)$. Here λ is the regularization parameter for TM regularization, and Z^{-1} is the normalization constant for the probability.

$$\mathcal{L}_{reg}(o) = \lambda \sum_{\mathbf{x}} |\nabla o(\mathbf{x})|^2 \tag{26}$$

where $\nabla o(\mathbf{x}) = (|\frac{\partial o(\mathbf{x})}{\partial x}|^2 + |\frac{\partial o(\mathbf{x})}{\partial y}|^2 + |\frac{\partial o(\mathbf{x})}{\partial z}|^2)$

Good's Prior This function was proposed by Good as a measure for discriminating slightly displaced density curves or hyper-surfaces. Verveer and Jovin [26] developed a superresolution restoration algorithm on this prior for fluorescence microscopy. It has the following form for the penalty functional:

$$\mathcal{L}_{reg}(o) = \sum_{\mathbf{x}} \frac{|\nabla o(\mathbf{x})|^2}{o(\mathbf{x})} \tag{27}$$

2.3.2 ℓ_1 regularization

If $(O = o)$ is a low-order, homogeneous, isotropic MRF, over a 6 member neighborhood $\eta_{\mathbf{x}} \in \eta$ of the site $\mathbf{x} \in \Omega$,

$$P[X = o(\mathbf{x})] \propto \frac{1}{Z_\lambda} e^{-\lambda \sum\limits_{\mathbf{x}} |\nabla o(\mathbf{x})|}, \tag{28}$$

$$|\nabla o(\mathbf{x})| = (\sum_{\mathbf{x}' \in \eta_{\mathbf{x}}} (o(\mathbf{x}) - o(\mathbf{x}'))^2)^{\frac{1}{2}},\ Z_\lambda = \sum_{o \in \mathcal{I}(\otimes)} e^{-\lambda \sum\limits_{\mathbf{x}} |\nabla o(\mathbf{x})|} \tag{29}$$

λ is called the hyperparameter, $|\nabla o(\mathbf{x})|$ is the *potential function* and Z_λ is a normalizing constant called the *partition function*. In order that Z_λ is finite we must restrict the possible values of $o(\mathbf{x})$ so that the numerical gradient of $\nabla o(\mathbf{x})$ is also bounded.
When used as a prior model for the object, we have to minimize the following functional:

$$\mathcal{L}_{reg}(o(\mathbf{x})) = \lambda \sum_{\mathbf{x}} |\nabla o(\mathbf{x})| \tag{30}$$

It has been observed that a ℓ_1 norm over ∇o rather than the ℓ_2 norm allows to impose edge preserving smoothing of the solution [28]. This is because the smoothing process introduced by the ℓ_1 norm works only in the direction tangential to the level lines (edges) and not in the orthogonal direction. Hence, the Total Variation (TV) regularization (Rudin *et al.* [34]) is able to preserve discontinuities better. TV has been mostly used for denoising 2D images with additive noise like characteristics. However, it was Charbonnier [35] who used a similar method to reconstruct 3D tomographic images in the presence of shot or poisson noise. A direct 3D extension of ROF's iterative noise removal algorithm is described by Persson *et al.* [36] for their work on 3D tomographic images.

2.3.3 $\ell_2 - \ell_1$ regularization

Like the ℓ_1 norm, Charbonnier et al. [37] introduced a possible solution to combine the advantages of the regularizations described above. These priors are called quadratic-linear or $\ell_2 - \ell_1$ priors. This method was adopted by Chenegros et al. [38] for deconvolution of 3D retinal images. This is summarised as follows:

$$\mathcal{L}_{reg}(o) = \lambda \delta^2 \sum_{\mathbf{x}} \phi(\frac{\nabla o(\mathbf{x})}{\delta}), \tag{31}$$

where $\phi(\cdot) = |\cdot| - \log(1 + |\cdot|)$ is a functional that is quadratic for small gradients and linear for large ones.

2.3.4 Entropy Prior

The entropy prior distribution of the object is given by,

$$P(o(\mathbf{x})) = Z^{-1} \exp(\lambda S(|Co(\mathbf{x})|, |Cm(\mathbf{x})|)), \tag{32}$$

where the absolute value $|\cdot|$ is applied to each element of $|Co|$ and $|Cm|$, corresponding to the assumption that the absolute values of linear combinations of pixels have an entropy distribution. However, because the true object is unknown, m is often chosen to be a constant image leading to a preference over smooth solutions. The entropy function S is defined by,

$$S(\mathbf{x}, \mathbf{y}) = \sum_{l=1}^{M} (\mathbf{x}_l - \mathbf{y}_l - \mathbf{x}_l \log(\frac{\mathbf{x}_l}{\mathbf{y}_l})), \tag{33}$$

where the model $\mathbf{y}$ reflects the prior knowledge of $\mathbf{x}$ and $\mathbf{x}_l$, $\mathbf{y}_l$ are the l^{th} elements of vectors $\mathbf{x}$ and $\mathbf{y}$. Thus, the energy functional is

$$\mathcal{L}_{reg}(o) = \lambda |Co(\mathbf{x})|^T \log(\frac{|Co(\mathbf{x})|}{e|Cm(\mathbf{x})|}). \tag{34}$$

In the absence of prior knowledge and for simplifying the numerical computations, we set $C = I$

2.3.5 Wavelet Priors

It was noticed that finer details of the 3-D captured image are more sensitive to noise, mismatches in the PSF, under sampling, etc. Wavelet decomposition can perform quite well by analyzing the different scales separately in the deconvolution process. In literature there are many papers that have carried out deconvolution in 2-D like [27] [39] [40] [41] [42]. However, development of a 3-D deconvolution scheme using wavelets is still unexplored because the inverse kernel is known only approximately, a situation routinely encountered in 3-D microscopy. An approximate PSF model makes it attractive to propose a computationally efficient wavelet deconvolution scheme dealing with huge 3-D images. However, wavelet denoising methods can act as an efficient regularization scheme for 3-D deconvolution, enabling higher resolution. These methods not only keep the frequency bandwidth of the final restored image nonincreasing, but also suppress many of the artifacts that would be present at a given step of deconvolution without denoising. Generally, in combination with more classical deconvolution algorithms, wavelet denoising methods can provide a robust and efficient deconvolution scheme [15].

2.4 PSF Models

The deconvolution algorithms discussed in Section 2.2 relies on an accurate model or knowledge of the PSF. The PSF can be defined either theoretically using a mathematical model of diffraction (see Section 2.4.2) or empirically by acquiring a 3D image of a fluorescent bead [43][44].

2.4.1 Experimental PSF

Because PSFs play an essential role in accurate reconstruction of the object intensity distribution, considerable effort has been directed at measuring and characterizing their properties [6] [45]. An empirically obtained PSF is an image of a small fluorescent bead positioned so as to approximate a point object of unit intensity positioned in the slide. To improve the statistics of the PSF data, several bead data sets are acquired and averaged. It is the averaged data that is used by the deconvolution algorithm. The experimental PSF is very difficult to obtain especially for fluorescence microscopes and they often change with slight deviations from the physical conditions. It is thus necessary to maintain the same conditions when obtaining the real images. Noise is always present in an experimentally obtained PSF and has to be handled before the deconvolution process. An empirical PSF can deviate significantly from perfect symmetry [17] [46]. This deviation or aberration, is often caused by mismatch in the refractive index between the specimen immersion and the objective medium.

2.4.2 Physical PSF models

In this section, we present a simple physical model of the PSF assuming only the diffraction-limited nature of the imaging system. The Optical Transfer Function for incoherent illumination can be obtained from knowledge of the PSF, and vice versa, by using the Fourier-transform relationship between them. P. A. Stokseth [6] obtained the OTF and PSF of an aberration-free defocused optical system from the corresponding pupil function. This model is useful to produce blur on synthetic objects and to test the deconvolution algorithms. The diffraction approximated confocal PSF is the square of the WFM PSF. If we assume that the excitation and emission wavelengths are the approximately the same as in the case of fluorescence, then the PSF is given as (after including the effect of the pinhole):

$$h(\mathbf{x}) = |A(r) * \mathcal{F}(P_{\lambda_{em}}(\mathbf{x}))|^2 \cdot |\mathcal{F}(P_{\lambda_{ex}}(\mathbf{x}))|^2 \tag{35}$$

where $A(r)$ models the pinhole (in 2D) that limits the out-of-focus light and P_λ is the pupil function given by [3] [6]:

$$P_\lambda(r,z) = \Pi_\rho(r) \cdot e^{i\frac{2\pi}{\lambda}W(z)} \tag{36}$$

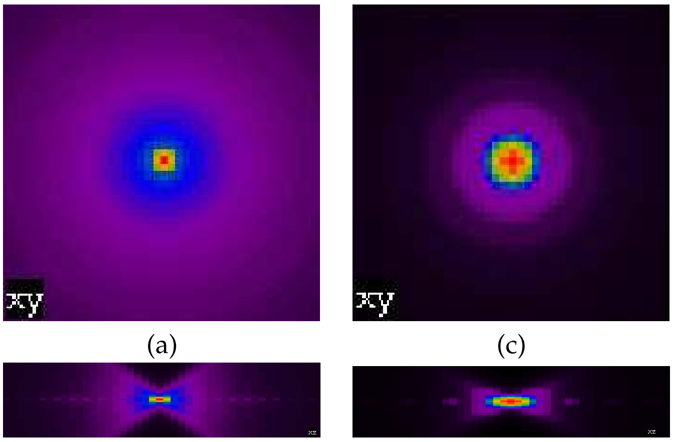
where,

$$W(z) = \frac{1}{2}z(1-\cos\alpha) \tag{37}$$

$\rho = \frac{NA}{\lambda}$, and α is the semi-aperture angle of the objective lens.
If $r^2 = x^2 + y^2$, and R is the radius of the pinhole, then

$$A(r) = \begin{cases} 1 & \text{if } r \leq R \\ 0 & \text{elsewhere} \end{cases} \tag{38}$$

As the pinhole size becomes infinitely small, $\lim\limits_{R \to 0} A(r) = \delta(r)$ and is a perfect pinhole. In reality, since the pinhole size cannot be infinitely small, a usable diameter of 1AU (see Appendix A) allows more than 70% of the in-focus light to the counter. Fig. 6 shows the radial and axial slices of analytically generated WFM and CLSM PSFs. This approximation is only valid when imaging thin samples ($< 5\mu m$ in thickness). When imaging thick biological tissues, 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 for all specimens. It has been shown [5] that when focusing into water using an oil immersion objective the signal level falls to 40% when focusing only $5mm$ beneath the surface (cover glass) and is below 10% at $15mm$. Thus restoration of thick samples or viewing deeper into the specimen becomes difficult with just the diffraction-limited approximation. Hence there is a need to model the PSF to include depth-dependent variations. Computed PSFs are noiseless, but do rely on a simplified model of the microscope and they often fall short on being able to reliably model the actual PSF. Furthermore, the assumed experimental parameters, which are used in the computational model, may differ from their actual values. For example, the PSF is very sensitive to some parameters like



(a) (c)

(b) (d)

Figure 6: (a) Radial and (b) axial PSF slices for a Wide-Field Microscope (size $4.55 \times 4.55 \times 19.38$ μm^3), and (c) (d) for Confocal Microscope (size $2.03 \times 2.03 \times 9.09$ μm^3) assuming an excitation wavelength of 488nm, N$A = 1.2$, $\mu_{specimen} = 1.515$, $\mu_{lens} = 1.515$

the immersion oil index, the depth of the specimen under the coverslip, or the numerical aperture of the objective.

2.4.3 Blind restoration

The problem of blind deconvolution is reduced to answering the following question:
"How does one estimate an original object and the PSF of the microscope, given only the observation data?"
Many methods use an iterative approach to estimate the PSF and the object with no prior information on the object (see for example the book [47]). However, unlike in other applications, the imaging of a specimen using fluorescence microscopes is often a non-repeatable process. For example, when imaging live cells, overexposure to the excitation laser might kill them due to abnormal levels of the radiation or "*Photobleaching*" wherein the fluorophore (here GFP) itself is destroyed. So given a single observation of the sample, it is very difficult to reconstruct the original specimen and the PSF from just this information. If the problem of deconvolution is ill-posed, that of Blind Deconvolution is even more so because we are increasing the number of unknowns without increasing the amount of data. Intuitively, if we forget about the effect of noise and consider the blurring observation model in the Fourier space: $\mathcal{F}(i) = \mathcal{F}(h) \cdot \mathcal{F}(o)$, and try to find both h and the object o solely from this product. Several solutions can respond to this problem, for example, if $(h,\, o)$ is a solution, then the trivial solution is that h is a dirac and $o = i$ or vice versa. If the blur function is irreducible, the problem admits several solutions for (h, o). A function h is said to be irreducible if it cannot be broken down as the convolution of two functions h_1 and h_2 (differing from Dirac). If h is not irreducible, 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 $(\alpha \cdot h,\, \frac{1}{\alpha} o)\ \forall \alpha > 0$ are solutions too. This ambiguity can be waived by imposing a forced normalization on h as in (40). Thus broadly speaking, a way of reducing the space of possible solutions and to regulate the problem is to introduce constraints on h and o.
It is for these reasons that there has been increased research activity recently for the blind restoration of images. Blind deconvolution (BD) methods simultaneously estimate the microscope PSF and the original 3-D specimen image. In optical microscopy, Holmes [48] was the first to propose a maximum likelihood-estimation-based blind deconvolution algorithm to deblur simulated images. The success of his algorithm was due to properly constraining the PSF estimate and the specimen especially in the simulation studies. While Markham and Conchello [49], presented their parametric model for the PSF and an estimation method utilizing this model. Our approach is to consider an a priori model on the object o (see Section 2.3) and a parametric model of the PSF h (see Section 3.1), and alternatively minimize a cost function, first w.r.t the object and then use the estimate to determine the PSF parameters. Without a model on the object or the PSF, direct minimization of the cost function can yield many possible solutions for o and h.

3 Joint estimation of the PSF and the specimen functions

3.1 3D Separable Gaussian PSF model

A regularization model on the PSF (similar to that on the specimen) has been suggested in some earlier works [48] [50]. However, there are three problems in applying a Total Variation (TV) [34] kind of regularization on the PSF. Firstly, it is not able to model the continuity and regularity in the PSF. Secondly, the recovered PSF will be very much dependent on the structure in the object/specimen [51]. Finally, the hyperparameter for such a model is highly dependent on the amount of defocus in the observed image. This means, that the parameter will vary drastically from one image sample to another and will have to be tuned manually for each restoration or estimated automatically.
The imperfections in an image formation system normally act as passive operations on the data, i.e. they do not absorb or generate energy. Thus, when an object goes out of focus it is blurred, but it's total intensity remains constant. Consequently, all energy arising from a specific point in the fluorescent specimen should be preserved, yielding:

$$\int_{\mathbf{x}} h(\mathbf{x})d\mathbf{x} = 1 \tag{39}$$

In the spatial domain, the discrete PSF is constrained to satisfy the following normalization:

$$\sum_{\mathbf{x}\in\Omega} h(\mathbf{x}) = 1.0 \tag{40}$$

and the follwing positivity constraints:

$$h(\mathbf{x}) \geq 0, \forall \mathbf{x} \in \Omega \tag{41}$$

It was shown by Santos and Young [52] (for the 2D case) and by Zhang et al. [53] (for the 3D case) that the diffraction-limited PSF is well modeled by a Gaussian function as:

$$\begin{gathered} h_{\sigma_r,\sigma_z}(r,z) = \frac{1}{Z_{\sigma_r,\sigma_z}} e^{\left(\frac{-r^2}{2\sigma_r^2} - \frac{z^2}{2\sigma_z^2}\right)}, \\ r = \sqrt{(x^2+y^2)},\ Z_{\sigma_r,\sigma_z} = (2\pi)^{\frac{3}{2}}|\Sigma|^{\frac{1}{2}} \end{gathered} \tag{42}$$

A diffraction-limited PSF has a circular symmetry about the z-axis and mirror symmetry about the central xy-plane. Thus, the covariance matrix should be diagonal [53] and $|\Sigma| = \sigma_r^4\sigma_z^2$. Then the *partition function* or the central '*surface brightness*' will be $Z_{\sigma_r,\sigma_z} = (2\pi)^{\frac{3}{2}}\sigma_r^2\sigma_z$. Consequently, an arbitrary image volume can be approximated as the convolution of a separable 3D Gaussian model and the actual volume of interest. This parameterized model is a good way to impose constraints on the PSF and regularize its estimation as the number of unknown parameters is reduced to just two viz. σ_r and σ_z.

3.2 Bayesian approach for joint object and blur parameter estimation

By the Bayes theorem, the posterior probability $P(X = o|Y = i)$ is:

$$P(X = o|Y = i) \propto \frac{1}{Z_\lambda} e^{-\lambda \sum\limits_{\mathbf{x}} |\nabla o(\mathbf{x})|} \cdot \prod_{\mathbf{x} \in \Omega} \frac{[h(\boldsymbol{\theta}) * o](\mathbf{x})^{i(\mathbf{x})} e^{-[h(\boldsymbol{\theta}) * o](\mathbf{x})}}{i(\mathbf{x})!} \tag{43}$$

where $\boldsymbol{\theta}$ are the parameters of the PSF model. The estimates for o and h (or $\boldsymbol{\theta}$) are obtained from the joint probability as:

$$\begin{aligned} (\hat{o}, \hat{h}) &= \arg\max_{(o,h)} P(X = o, H = h|Y = i) \\ &= \arg\min_{(o,h)} (-\log[P(X = o, H = h|Y = i)]) \end{aligned} \tag{44}$$

Thus, the cost function to be minimized is:

$$\mathcal{L}(o, h) = \lambda \sum_{\mathbf{x} \in \Omega} |\nabla o(\mathbf{x})| - \sum_{\mathbf{x} \in \Omega} i(\mathbf{x}) \log[h(\boldsymbol{\theta}) * o](\mathbf{x}) + \log[Z_\lambda] + \sum_{\mathbf{x} \in \Omega} [h(\boldsymbol{\theta}) * o](\mathbf{x}) \tag{45}$$

(45) can be rewritten in terms of the data fidelity term and the regularization term as:

$$\mathcal{L}(o|\hat{\boldsymbol{\theta}}) = \mathcal{L}_{obs}(o) + \lambda \sum_{\mathbf{x} \in \Omega} |\nabla o(\mathbf{x})| = \mathcal{L}_{obs}(o) + \mathcal{L}_{reg}(o) \tag{46}$$

Note that $\mathcal{L}_{obs}(o)$, the data fidelity term is convex and if h is positive definite, then $\mathcal{L}_{obs}(o)$ is strictly convex in o.

3.2.1 Estimation of the object

Given the estimated value (or initial) of $\hat{\boldsymbol{\theta}}$, and as Z_λ is not dependent on h or o, $\mathcal{L}(o, h)$ has now the form:

$$\mathcal{L}(o|\hat{\boldsymbol{\theta}}) = \sum_{\mathbf{x} \in \Omega} [h(\hat{\boldsymbol{\theta}}) * o](\mathbf{x}) - \sum_{\mathbf{x} \in \Omega} (i(\mathbf{x}) \log[h(\hat{\boldsymbol{\theta}}) * o](\mathbf{x})) + \lambda \sum_{\mathbf{x} \in \Omega} |\nabla o(\mathbf{x})| \tag{47}$$

We now use the Maximum A Posteriori (MAP) algorithm (that is equivalent to the penalized Maximum Likelihood (ML) [54]), with a prior distribution on the object o.

Richardson-Lucy (RL) algorithm with TV regularization The Euler-Lagrange equation for minimizing $\mathcal{L}(o|\hat{\boldsymbol{\theta}})$ in (46) w.r.t θ is:

$$1 - h(-\mathbf{x})(\frac{i(\mathbf{x})}{(h * o)(\mathbf{x})}) - \lambda\, div(\frac{\nabla o(\mathbf{x})}{|\nabla o(\mathbf{x})|}) = 0 \tag{48}$$

where div stands for the divergence (see [28] for details). This can be solved for the object o by adopting a RL algorithm ([55] [56]) with TV regularization. This is an Expectation-Maximization (EM) algorithm [57] for computing the Maximum a Posteriori estimate.
We adopt an explicit scheme that proposes to minimize (46) by the following multiplicative algorithm:

$$o_{n+1}(\mathbf{x}) = [\frac{i(\mathbf{x})}{(o_n * h_{\hat{\boldsymbol{\theta}}}(\mathbf{x}))} * h_{\hat{\boldsymbol{\theta}}}(-\mathbf{x})] \cdot \frac{o_n(\mathbf{x})}{1 - \lambda div(\frac{\nabla o_n(\mathbf{x})}{|\nabla o_n(\mathbf{x})|})} \tag{49}$$

where $(\cdot)$ denotes the Hadamard multiplication (component wise) and n the iteration number.
To stop the iterations, we define the difference measure between two successive iterations. If the difference is smaller than a threshold (t), we stop the algorithm to get the final deconvolved image. The criterion is defined as:

$$\chi_{n+1} = \frac{\sum\limits_{\mathbf{x}\in\Omega} |o_{n+1}(\mathbf{x}) - o_n(\mathbf{x})|}{\sum\limits_{\mathbf{x}\in\Omega} o_n(\mathbf{x})} < t \tag{50}$$

The typical values chosen for t are 10^{-4} or 10^{-6}, depending on the precision desired.

3.2.2 Parameter estimation on complete data

Once the object o is estimated by using the iterative algorithm that is described above, the PSF can be obtained by minimizing the following function:

$$\mathcal{L}(h|i, \hat{o}) = -\sum_{\mathbf{x}\in\Omega}(i(\mathbf{x}) \log[h * \hat{o}](\mathbf{x})) + \sum_{\mathbf{x}\in\Omega}[h * \hat{o}](\mathbf{x}) \tag{51}$$

The terms that are independent of h have been excluded. A direct, closed-form solution does not yield the exact solution in this case. However, by introducing constraints on the PSF, we greatly reduce the number of free parameters to describe it and the estimation is realizable as described below.
The log-likelihood is symmetric about both the object to be estimated and the point-spread function. The PSF can be estimated by the Shepp-Vardi solution [58], which is based on ML and is similar to the RL algorithm described previously for estimating o as:

$$\hat{h}_{k+1}(\mathbf{x}) = \hat{h}_k(\mathbf{x}) \sum_{\mathbf{x}\in\Omega} [\frac{\hat{o}(-\mathbf{x})}{[h * \hat{o}](\mathbf{x})}] i(\mathbf{x}) \tag{52}$$

There is however no constraint imposed on the PSF and the solution is thus not always unique. One way of introducing a constraint on the PSF and of avoiding the other possible solutions, is to limit the bandwidth of the Point-Spread function by parameterizing it. This

was introduced in Section 3.1.
The method outlined in Section 3.2.1 requires the knowledge of parameters $\boldsymbol{\theta} = (\sigma_r, \sigma_z)$. However, due to the invariance property of ML estimation, we can say:

$$\hat{h}_{ML}(\mathbf{x}) = h(\mathbf{x}, \hat{\boldsymbol{\theta}}_{ML}) \tag{53}$$

is the MLE of the PSF.
From (51), we get:

$$\mathcal{L}(\boldsymbol{\theta}|i,\, \hat{o}) = -\sum_{\mathbf{x}\in\Omega}(i(\mathbf{x})\log[h(\boldsymbol{\theta}) * \hat{o}](\mathbf{x})) + \sum_{\mathbf{x}\in\Omega}[h(\boldsymbol{\theta}) * \hat{o}](\mathbf{x}) \tag{54}$$

The parameters of the PSF can be estimated as (see Appendix C),

$$\hat{\boldsymbol{\theta}}^{(k+1)} = \hat{\boldsymbol{\theta}}^{(k)} - \alpha^{(k)}\nabla_{\boldsymbol{\theta}}\mathcal{L}(\hat{\boldsymbol{\theta}}^{(k)}) \tag{55}$$

where $\alpha^{(k)}$ and $\nabla_{\boldsymbol{\theta}}\mathcal{L}(\hat{\boldsymbol{\theta}}^{(k)})$ are the step size and the search direction at iteration k respectively. The search direction involves calculating $\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})$, w.r.t the PSF parameters:

$$\begin{aligned}\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}) = \sum_{\mathbf{x}\in\Omega}((h_{\boldsymbol{\theta}} * \hat{o})(\mathbf{x}) - \frac{i(\mathbf{x})}{(h * \hat{o})(\mathbf{x})} \cdot (h_{\boldsymbol{\theta}} * \hat{o})(\mathbf{x}))\\ \text{where, } h_{\boldsymbol{\theta}} = \frac{\partial}{\partial\boldsymbol{\theta}}h(\boldsymbol{\theta});\ \boldsymbol{\theta} \in \{\sigma_r, \sigma_z\}\end{aligned} \tag{56}$$

We stop the computation if the difference measure between two successive iterations is smaller than a specific threshold ϵ (in practice 10^{-3} or 10^{-4}), and assume the last estimation as the best one.

$$\chi_{k+1} = \frac{|\hat{\boldsymbol{\theta}}^{(k+1)} - \hat{\boldsymbol{\theta}}^{(k)}|}{\hat{\boldsymbol{\theta}}^{(k)}} < \epsilon \tag{57}$$

Alternate Minimization algorithm

The pseudo-code of the algorithm that was used for the joint estimation of the parameters of the PSF model and the specimen function is as given below:

```
1.  Calculate the permissible theoretical range $[\sigma_{r,min},\, \sigma_{r,max}]$
equations (69, 67) and $[\sigma_{z,min},\, \sigma_{z,max}]$ equations (71, 70).

2.  The hyperparameter, $\lambda$ is fixed at an optimal value of $0.002$ and
the threshold $t$ (50) is chosen as $10^{-4}$ for the deconvolution
algorithm (RL+TV) [13].
```

3. Assign $k = 0$, and choose the theoretical parameters as $\hat{\sigma}_{r,\,k}$ and $\hat{\sigma}_{z,\,k}$.
4. Calculate $\hat{h}_k$ from the values of $\hat{\sigma}_{r,k}$ and $\hat{\sigma}_{z,k}$.
5. The degraded image of the object $i(\mathbf{x})$ is chosen as the initial object estimate $\hat{o}_k(\mathbf{x})$.
6. Deconvolve $\widehat{o}_k(\mathbf{x})$ with the RL+TV algorithm (49) to get a new estimate of the specimen function $\hat{o}_{k+1}(\mathbf{x})$ when the error between successive iteration estimates do not differ by t.
7. The parameters of the PSF are estimated using the updated specimen function estimate obtained in the previous step. The ***Conjugate-Gradient (CG) algorithm*** (83) that minimizes the objective function $\mathcal{L}(\sigma_r, \sigma_z | \hat{o}(\mathbf{x}), \lambda)$ gives the values $\hat{\sigma}_{r,k+1}$ and $\hat{\sigma}_{z,k+1}$.
8. k is incremented and the error ($\chi_k = \frac{(\hat{\theta}(k) - \hat{\theta}(k-1))}{\hat{\theta}(k)}$) calculated. The result of this is used in step 6 and the AM algorithm is continued until $\chi_k \leq \epsilon$.

4 Results

In this section, we present the outcome of our alternate minimization method on some synthetic and real data.

4.1 Experiments on Simulated Data

The 3D simulated test object (see Fig. 7(a)) used is of dimension $128 \times 128 \times 64$, with XY and Z pixel sizes of $50nm$ and $230nm$ respectively. The observed data was then generated by using an analytical model of the microscope PSF [59] (with a pinhole diameter of 1 AU), and the noise is mainly shot noise due to low-photon imagery, modelled as Poisson statistics (see Fig. 7(b)). We present here the deconvolution results using the Richardson-Lucy multiplicative algorithm regularized using a functional derived from the Total Variation. The deconvolution and the estimation results are as shown in Fig. 7(c), (d) respectively.

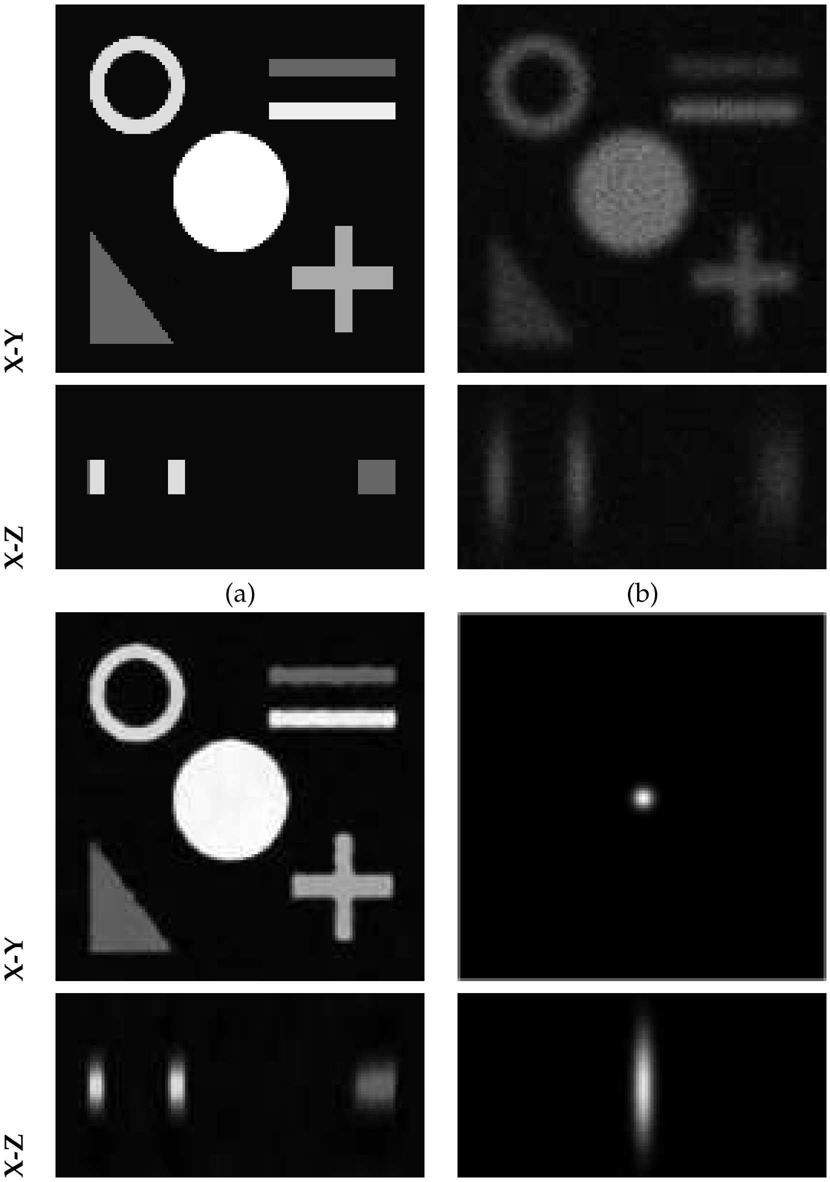
If the synthetic object o is assumed to be known a priori, then estimation of the true parameters of the PSF is straight forward as the cost function is convex in the neighborhood of optimal σ_r as illustrated in Fig. 8.

The optimal stopping threshold t (50), or the hyperparameter λ, dictates the amount of reduction in the blur and noise-edge effects respectively. Correct estimation and convergence of the parameters of the PSF model is dependent on the stable estimate of the object. For the simulated data case, the optimal hyperparameter λ was chosen to be 0.002, and the stopping threshold t for two successive iterations as 10^{-4}. Fig. 9 shows the reduction in the cost function with iterations of the CG algorithm and the approach of the estimated radial spread to the stable value when the estimate of the object is assumed to be known.

Since there are no constraints on the PSF, if the parameters are initialized to small values (PSF is Dirac in nature), the growth of the PSF to the desired size is not guaranteed. This point is illustrated in Fig. 10 where the dark blue regions are the area of local minima. To avoid this problem, we choose the initial parameters to be larger than the theoretically calculated values [53] (see Appendix B) and descend to the true value (as shown by the white arrow heads). If the microscope is assumed to have a NA of 1.4, excitation wavelength peak of 488nm, emission spectra peak at 520nm and refractive index of objective immersion medium 1.515, then the non-paraxial parameters are 75.734nm (σ_r) and 150.066nm (σ_z), and paraxial (if the light ray is close to and nearly parallel with the optical axis) parameters are 76.551nm (σ_r) and 230.841nm (σ_z) respectively.

Quantitative Analysis

The *I-divergence* or generalized Kullback distance is used for measuring the quality of the restored image in comparison to the original image. The I-divergence, $I(u, v)$, between two non-negative distributions u and v measures the discrepancy or the distance of the function



(a) (b)

(c) (d)

Figure 7: (a) Composite synthetic object, (b) observed image with the analytical blur model [6] and Poisson noise, (c) image restored after RL+TV deconvolution using the estimated PSF, (d) estimated microscope PSF.

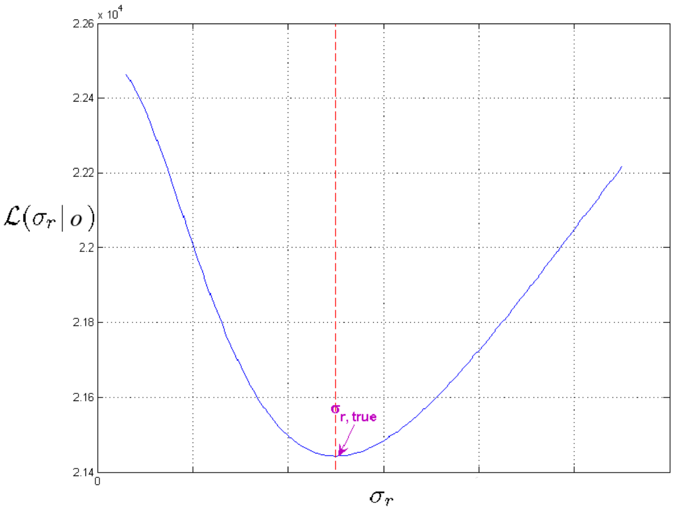


Figure 8: Cost function as a property of the unknown PSF parameter when the initial object and observation are known.

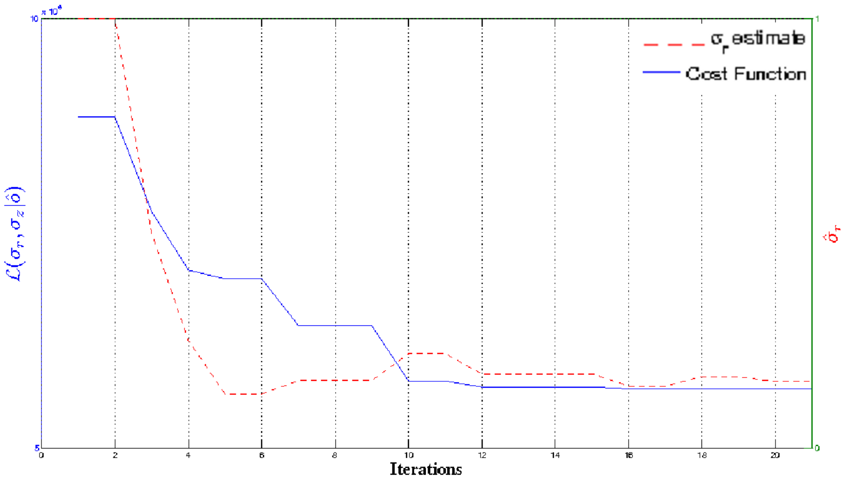


Figure 9: Convergence of the cost function and radial parameter by the conjugate gradient method (when the original object is known). The Y axis is left-scaled for the cost function $\mathcal{L}(\sigma_r, \sigma_z | \hat{o})$ and right-scaled for the PSF parameter respectively.

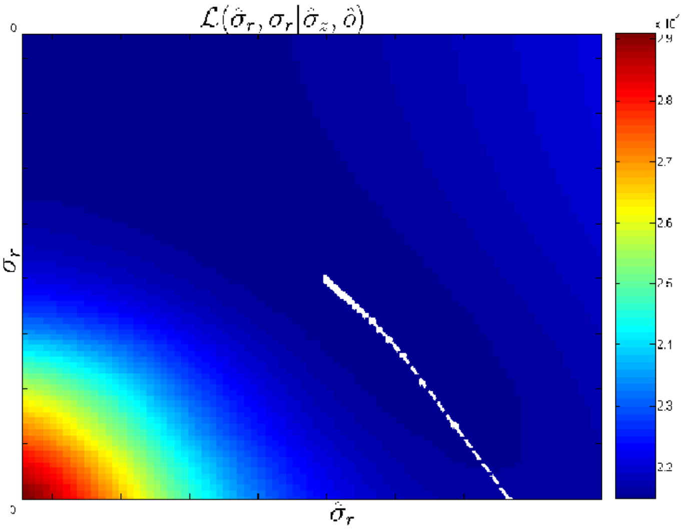


Figure 10: Plot of the energy as a function of the true and the estimated parameter values. White arrows points the direction followed by the estimation algorithm.

v to u. It was proved to be consistent [60] with the *Csiszar's* axioms [61].

$$I(u, v) = \sum_{\mathbf{x} \in S} u(\mathbf{x}) \log \frac{u(\mathbf{x})}{v(\mathbf{x})} - \sum_{\mathbf{x} \in S} [u(\mathbf{x}) - v(\mathbf{x})] \tag{58}$$

After 7 *joint* iterations of the alternated minimization algorithm, when the stopping criterion (57) is reached, the final *I-divergence* was 0.767 and there was 43.81% improvement over the standard RL algorithm. It was also observed that the theoretically calculated and the experimentally estimated PSF parameter values differ by about 16% and 14.5% for the radial and the axial case respectively. Fig. 11 compares the estimated 3D PSF with the analytically modeled [6] PSF along the one direction of the radial plane. The error in the estimation given by the ℓ^2 residual of the PSF is $< 0.2\%$.

4.2 Experiments on Real Data

4.2.1 Imaging Setup

The Zeiss LSM 510 confocal microscope mounted on a motorized inverted stand (Zeiss Axiovert 200M) equipped with an *ArKr* laser of wavelength of 488nm is the excitation source, and Band Pass (BP) filter transmits emitted light within the band $505 - 550$nm. The objective lens arrangement is a Plan-Neofluar with 40X magnification having a Numerical Aperture of 1.3 and immersed in oil. The pinhole size was fixed at 67μm.

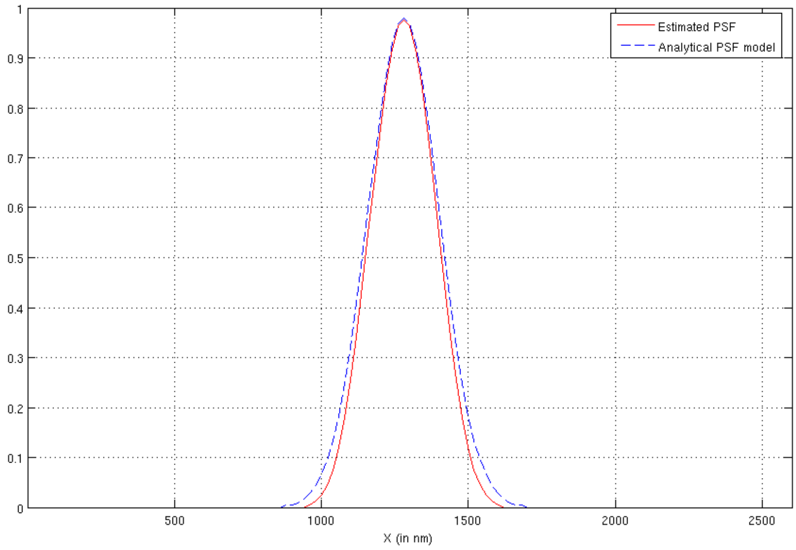


Figure 11: Comparison of the estimated (continuous line) and analytically computed (dotted line) PSF for a CLSM.

4.2.2 Specimen description

The specimen that was chosen for the experiments is the embryo of a *Drosophila melanogaster* (see Fig. 12). It was mounted and tagged with the Green Fluorescence Protein (GFP). This setup is used for studying the sealing of the epithelial sheets (*Dorsal Closure*) midway during the *embryogenesis*. The images (© Institute of Signaling, Development Biolody & Cancer) were acquired with a XY pixel size of 50nm and a Z step size of 170nm, and the size of the volume imaged is $25.59 \times 25.59 \times 2.55\mu$m.

4.2.3 Deconvolution Results

The maximum intensity projection of the observed and restored data is shown in Fig. 12. Since the observation data had a high noise content, the regularization hyperparameter λ was set to a large value of 0.05. The deconvolution algorithm was stopped when the difference between subsequent estimates was lower than $t = 0.002$. The *alternate minimization* algorithm converged after 40 iterations of the joint RL-TV and CG algorithm. The PSF parameters were initialized to 300nm and 600nm for the radial and the axial case respectively, and the conjugate-gradient algorithm estimated them to be 257.9 and 477.9nm [62]. These are much larger than their corresponding theoretically expected values given in Section 4.1 [53]. It was verified that the proposed algorithm can not only estimate the actual

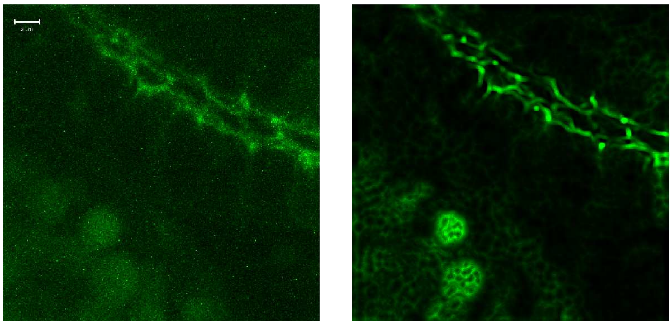
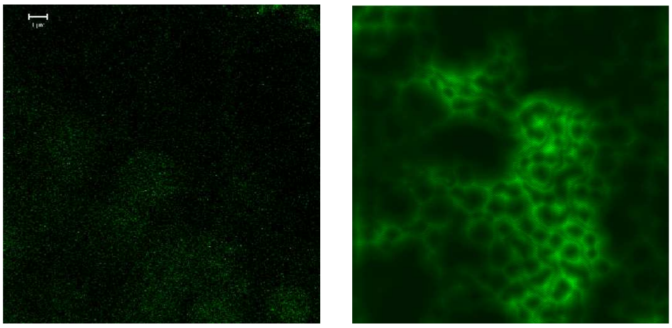


Figure 12: Maximum Intensity Projection of the original specimen (left) (© Institute of Signaling, Developmental Biology & Cancer UMR6543/CNRS/UNSA), and restored image (right) (© Ariana-INRIA/I3S). The intensity is scaled to $[0\,130]$ for display and white bar is $2\mu m$.

PSF from the experiments on synthetic data [62], but also provide much better deconvolution results in comparison to theoretical microscope PSF's (generated using the microscope settings).

It was noticed that in each slice of the chosen observation data, there was a significant amount of signal contributions from the neighboring slices too. A subjective analysis of the deconvolution results [63] showed that although (see Fig. 14) some structures and cellular walls are visible at 0μm depth of the original data, they are really absent at this level and appears only at 0.34μm in the restoration. The algorithm provides reasonably good deconvolution results even at depths of more than 1μm as shown in Fig. 13.



(a) (b)

Figure 13: Deconvolution results at a depth of $1.53\mu m$ (white bar is 1μm). (a) Observed image section (© Institute of Signaling, Developmental Biology & Cancer UMR6543/CNRS/UNSA), (b) Restored image (© Ariana-INRIA/I3S).

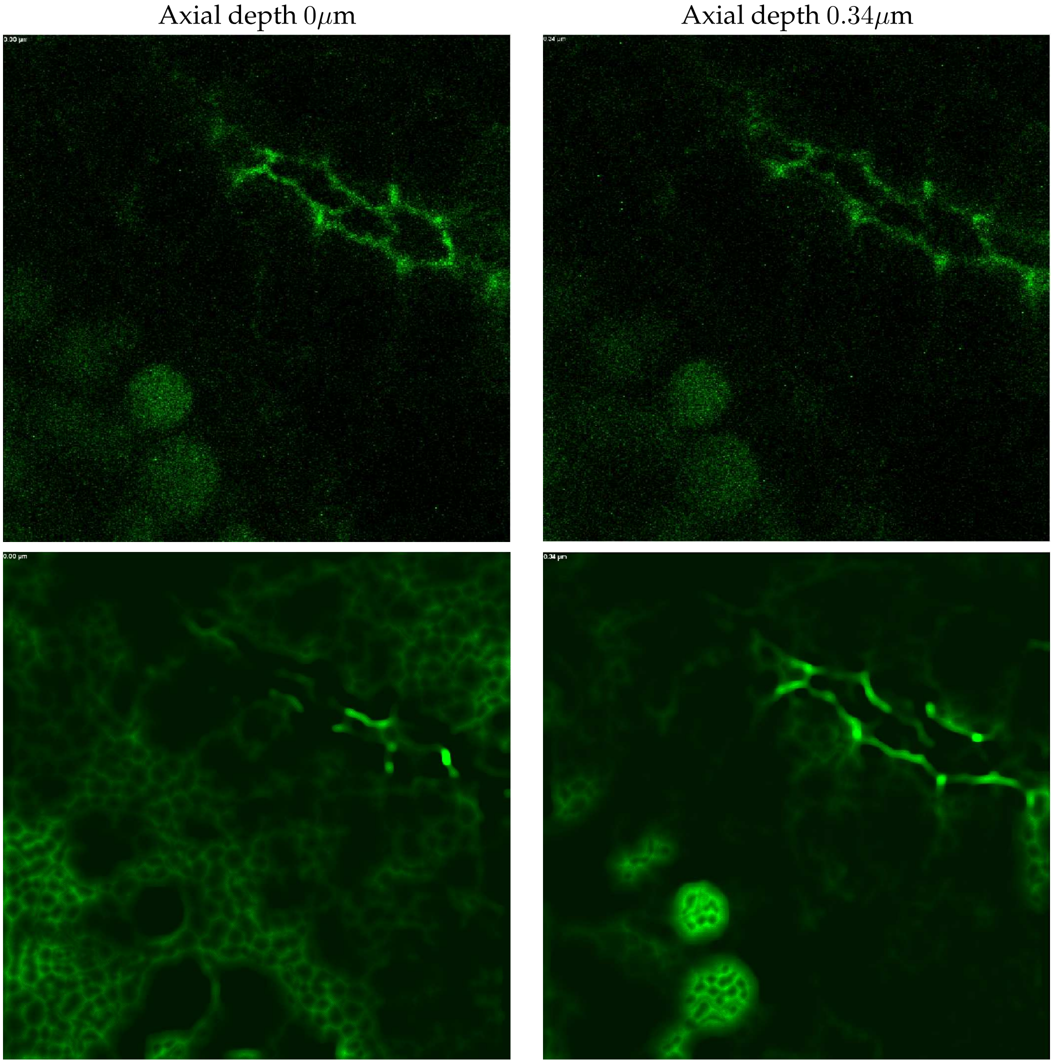


Figure 14: The observed image slices (top) (© Institute of Signaling, Developmental Biology & Cancer UMR6543/CNRS/UNSA) and restoration results (bottom) (© Ariana-INRIA/I3S) are shown at different depths. The intensity is scaled to $[0\,100]$ for display.

5 Conclusions and future work

In this report we have proposed an "*Alternate Minimization (AM) algorithm*" for the joint estimation of the microscope PSF and the specimen source distribution for a CLSM. We choose the RL algorithm for the deconvolution process as it is best suited for the Poisson data, and TV as the regularization model. A separable 3D Gaussian model best describes the PSF, and is chosen as the a priori model. We have experimented on real and simulated data, and the method gives very good deconvolution results and a PSF estimation close to the true value [62]. However, it should be noted that, all the out-of-focus light cannot be rejected and some noticeable haze and axial smearing remains in the images. The TV regularization preserves the borders better than the ℓ_2 norm, and the specimen is thought to be piecewise constant with discontinuities. However, small structures close to noise are not well restored (staircase effect), and some corners are rounded. A way to minimize the deblurring artifacts may be to estimate the hyperparameter of the regularization model [64] [65].

Restoration of thick samples or viewing deeper into the specimen becomes difficult with just the diffraction-limited approximation. Hence there is a need to model the PSF to include depth-dependent variations. However, in such as case, the classical Linear-Shift-Invariant (LSI) approximation is no longer valid. Future work is aimed at extending this algorithm to restore spherically aberrated (an optical effect occurring when the oblique rays entering a lens are focused in a different location than the central rays) observation data and also to improve the representation of the object to match the physical conditions.

A Glossary

Light microscopy terminologies

Numerical Aperture (NA)

The numerical aperture (NA) of a microscope objective lens is the half angle of the maximum ray of light entering it (according to Snell's law).

$$\mathrm{NA} = \mu_o \sin(\alpha) \tag{59}$$

where μ_o is the refractive index of the objective immersion medium and α is the semi-aperture angle of the objective lens

Lateral Resolution

The lateral resolution of a microscope is the ability of the microscope to discern 2 incoherent point sources separated by a small distance d in the specimen plane. The images are said to be resolved if $d \geq r_{Airy}$. The image of an infinitely small luminous object point is in itself not infinitely small (see Section 1.2.1), but is a circular Airy diffraction image with a small central bright disk and progressively weaker concentric dark and bright rings. The radius r_{Airy} of the first dark ring around the central disk of the Airy diffraction image depends on the emission wavelength λ_{em} in vacuum and the NA of the objective:

$$r_{Airy} = 0.61 \frac{\lambda_{em}}{\mathrm{NA}} \tag{60}$$

This is also known as the *Rayleigh criterion* and it relies on the assumption that the 2 point sources radiate incoherently. The factor 0.61 is the distance of the first minimum of an Airy function.

Two components of equal intensity should be considered to be just resolved when the principal intensity maximum of one coincides with the first intensity minimum of the other.

Axial Resolution

For defining the axial (z-axis) resolution, measured along the optical axis of the microscope, it is customary to use the 3-D diffraction image of a point source that is formed near the focal plane. It is defined as the minimum distance that the diffraction images of 2 points can approach each other along the axis of the microscope and yet be seen as two. For a widefield microscope, the axial resolution is related to the *depth of focus* of the microscope. For high numerical aperture, and no paraxial approximation,

$$R_z = 0.885 \frac{\lambda_{em}}{\mu_o - \sqrt{\mu_o^2 - \mathrm{NA}^2}} \tag{61}$$

Sampling Density

The sampling density is the number of recorded voxels per unit volume after converting to the digital mode. Thus, the larger the sample size (or sampling distance, the size of one voxel), the smaller is the sampling density.

Nyquist sampling

The Nyquist criterion determines the minimal sampling density needed to capture all the information of the specimen by the microscope into the image. When the sampling distance is larger than the Nyquist distance, information about the image is lost. This *Under Sampling* condition also gives rise to *Aliasing Artefact*. *Aliasing Artefact* may show up as jagged edges (staircasing) or fringes that are very hard to remove from the image.

Fluorescent Proteins

Fluorescent proteins are biological markers that can attach themselves to a protein of interest in the biological cell without impacting its functions.

CLSM terminologies

Backprojected pinhole radius

Backprojected pinhole is the microscope pinhole as it appears in the specimen plane. Its value is given by the physical pinhole size r_o and the total magnification of the detection system. The total magnification is the product of the (variable) objective magnification and a fixed internal magnification. Thus,

$$r_b \; = \; \frac{r_o}{m_o \, m_{sys}} \tag{62}$$

where m_o is the magnification factor of the objective and m_{sys} is the fixed internal magnification of the system (3.33 for Zeiss LSM 510 Confocal Microscope). In some microscopes with pinhole shapes other than circular, geometrical corrections may also be necessary.

Airy Units

Most confocal microscopes measure the pinhole size (diameter) with the *Airy Disk* (diameter) as unit. The backprojected pinhole radius can then be computed with:

$$r_b \; = \; N_{Ad} \, 0.61 \, \frac{\lambda_{ex}}{\mathrm{N}A} \tag{63}$$

with N_{Ad} the number of Airy disks, NA the Numerical Aperture of the lens, and $\lambda_{\mathrm{e}x}$ the excitation wavelength. In principle using $\lambda_{\mathrm{e}x}$ is not correct because the Airy diffraction pattern is formed by the emitted light. However, the excitation wavelength is used because it is better defined (is monochromatic and has no spectrum like the emission wavelength) and does not depend on settings of the adjustable band filters. Some microscopes use different pinholes for different channels, to really take one Airy disk (because the actual size of the Airy disk depends on the wave length of the radiation!). But other microscopes use just one pinhole, the same one for all channels. In this case, it is not longer true that exactly one Airy disk is taken for all the channels. Therefore to calculate a theoretical PSF for doing deconvolution, we need to know for what precise wave length is the pinhole designed. Moreover, the Refractive Index of the optical media is also relevant, because it changes the wavelength, so it is necessary to know if the manufacturer has taken this into account when giving the pinhole size. At the end, using Airy disk units is not such a good idea, because it introduces a lot of uncertainty on what is really going on inside the microscope.

Other terminologies

Cliques

A *clique* is a set of elements in the lattice of which all members are neighbors.

Anisotropy

Anisotropy (the opposite of *isotropy*) is the property of being directionally dependent. Something which is anisotropic, may appear different, or have different characteristics in different directions. A spherically symmetric Point Spread Function would be equal in all directions, and would produce isotropic smoothing of the images. Images from 3D Microscopes are anisotropic because the Point Spread Function (in ideal conditions) is cylindrically symmetric, larger along the optical axis.

Full Width at Half Maximum (FWHM)

The Half-Intensity Width (HIW) (also known as the Full Width at Half Maximum (FWHM)) of a curved function is the distance between the points where the intensity is half of the maximum one. The HIW of a Point Spread Function is a good estimation of the device Spatial Resolution, based on the Rayleigh Criterion. For Gaussian distributed data, $FWHM = 2\sqrt{2ln(2)}\sigma \sim 2.35\sigma$, where σ^2 is the variance in the data.

B Theoretical parameter limits

For faster convergence of the estimation algorithm, we calculate the theoretical Gaussian parameter limits for the confocal case [53]. If the PSF models are supposed to be diffraction-limited and aberrations ignored, based on the *Debye diffraction integral* the intensity distribution for a non-paraxial (NA > 0.7) objective can be considered as,

$$P(\mathbf{x}, \lambda) = C_0 \int_0^{\alpha} \sqrt{\cos\phi} J_o(kr\sin\phi) e^{-ikz\cos\phi} \sin\phi \, d\phi \tag{64}$$

where, $r = \sqrt{x^2 + y^2}$, α is the maximal convergence semi-angle of the objective, μ is the refractive index, $\sqrt{\cos\phi}$ is an apodization term, $k = \mu \cdot \frac{2\pi}{\lambda}$ is the wave number, and J_0 is the zero-order bessel function of first kind.
However, for paraxial objectives, the pupil function (assuming $\sin\phi \approx \phi$) can be written as

$$P(\mathbf{x}, \lambda) = C_1 e^{-ikz} \int_0^1 J_0(krt\sin\alpha) e^{\frac{i}{2}kzt^2\sin^2\alpha} t \, dt \tag{65}$$

where $C_1 = \alpha^2 C_0$, and $J_0(\cdot)$ is the Bessel function of first kind with order zero given by,

$$J_0(x) = \frac{1}{\pi} \int_0^{\pi} \cos(x\sin\theta) d\theta, \forall x \in \mathbb{R} \tag{66}$$

The lateral Gaussian parameter for 3-D PSFs, for the non-paraxial case (*i.e.* NA < 0.7) can be computed as [53]:

$$\hat{\sigma}_r^* = \sqrt{2}[\frac{2\sigma_{em,r}^4(e^{\frac{r^2}{2\sigma_{em,r}^2}} - 1) + r^2\sigma_{ex,r}^2}{\sigma_{ex,r}^2\sigma_{em,r}^4(e^{\frac{r^2}{2\sigma_{em,r}}} - 1)}]^{-\frac{1}{2}} \tag{67}$$

where

$$\sigma_{em,r} = \frac{1}{\mu k_{em}}[\frac{4 - 7\cos^{\frac{3}{2}}\alpha + 3\cos^{\frac{7}{2}}\alpha}{7(1 - cos^{\frac{3}{2}}\alpha)}]^{-\frac{1}{2}} \tag{68}$$

and $k_{em} = \frac{2\pi}{\lambda_{em}}$, $r = \frac{D}{2} = 0.61\frac{\lambda_{ex}}{NA}$. For the paraxial case (NA > 0.7), the lateral Gaussian parameter for 3-D PSF is:

$$\hat{\sigma}_r^* = \sqrt{2}[\frac{c_1^2}{r^2} + \frac{4c_2 J_0(c_2) J_1(c_2) - 8J_1^2(c_2)}{r^2(J_0^2(c_2) + J_1^2(c_2) - 1)}]^{-\frac{1}{2}} \tag{69}$$

where, $c_1 = k_{em} rNA$ and $c_2 = k_{em} rNA$.
Similarly, the 3-D axial Gaussian parameters for the PSF is (paraxial case),

$$\sigma_z^* = 2\sqrt{6}[\frac{c_1^2\mathrm{NA}^2}{r^2\mu^2} - \frac{48c_2^2[J_0^2(c_2) + J_1^2(c_2)] - 192J_1^2(c_2)}{\mu^2 k_{em}^2 r^4[J_0^2(c_2) + J_1^2(c_2) - 1]}]^{-\frac{1}{2}} \tag{70}$$

and for the non-paraxial case,

$$\sigma_z^* = \frac{\sigma_{ex,z}\sigma_{em,z}}{[\sigma_{ex,z}^2 + \sigma_{em,z}^2]^{\frac{1}{2}}} \tag{71}$$

where

$$\sigma_{em,z} = \frac{5\sqrt{7}(1-\cos^{\frac{3}{2}}\alpha)}{\sqrt{6}\mu k_{em}[4\cos^5\alpha - 25\cos^{\frac{7}{2}}\alpha + 42\cos^{\frac{5}{2}}\alpha - 25\cos^{\frac{3}{2}}\alpha + 4]^{\frac{1}{2}}} \tag{72}$$

C Algorithms in detail

In this appendix, we present the algorithms that are used in this research report in more detail. In the first half we present the discretized implementation of the TV regularization and towards the end the Conjugate-Gradient (CG) algorithm for parameter estimation.

Total Variation (TV) regularization

We know that the standard RL does not always converge to a suitable solution. That is because there is no information on the object. If we propose a prior model on the object, we regularize the solution. We will now maximize the a posteriori probability instead of maximizing the likelihood probability. The Total Variation norm (TV) was first introduced by Rudin *et al.* [34] as an iterative method of denoising an image. Since the TV has a ℓ^1 norm, unlike ℓ^2 norm of Tikhonov [32] and Miller, it's effects are to smooth homogeneous regions while preserving edges.
The 2D TV semi-norm is given by [34]:

$$TV(o(\mathbf{x})) = \int_{\mathbf{x}\in\Omega} |\nabla o(\mathbf{x})| d\mathbf{x} = \int \sqrt{\frac{\partial o^2}{\partial x} + \frac{\partial o^2}{\partial y}} dxdy \tag{73}$$

The discrete form of this equation can be written as:

$$\sum_{l,m} \sqrt{(o_{l+1,m} - o_{k,m})^2 + (o_{l,m+1} - o_{k,m})^2 + \eta^2} \tag{74}$$

An arbitrary parameter $\eta < 10^{-3}$ is introduced to ensure that the norm is differentiable at points where $(o_{l+1,m} - o_{k,m})^2 + (o_{l,m+1} - o_{k,m})^2 = 0$.
Our 3D implementation of the TV algorithm is described in more detail below. The discrete form of $div(\frac{\nabla o}{|\nabla o|})$ for a stable numerical scheme is the following [34] [13]:

$$\begin{aligned} div(\frac{\nabla o}{|\nabla o|}) = & \frac{1}{h_x}\Delta^x_- \frac{\Delta^x_+ o_{ijk}}{\sqrt{(\Delta^x_+ o_{ijk})^2 + m(\Delta^y_+ o_{ijk}, \Delta^y_- o_{ijk})^2 + m(\Delta^z_+ o_{ijk}, \Delta^z_- o_{ijk})^2}} \\ & + \frac{1}{h_y}\Delta^y_- \frac{\Delta^y_+ o_{ijk}}{\sqrt{(\Delta^y_+ o_{ijk})^2 + m(\Delta^z_+ o_{ijk}, \Delta^z_- o_{ijk})^2 + m(\Delta^x_+ o_{ijk}, \Delta^x_- o_{ijk})^2}} \\ & + \frac{1}{h_z}\Delta^z_- \frac{\Delta^z_+ o_{ijk}}{\sqrt{(\Delta^z_+ o_{ijk})^2 + m(\Delta^x_+ o_{ijk}, \Delta^x_- o_{ijk})^2 + m(\Delta^y_+ o_{ijk}, \Delta^y_- o_{ijk})^2}} \end{aligned} \tag{75}$$

defining the derivation as:

$$\begin{aligned}
\Delta_+^x o_{ijk} &= \frac{1}{h_x}(o_{(i+1)jk} - o_{ijk}) \\
\Delta_-^x o_{ijk} &= \frac{1}{h_x}(o_{ijk} - o_{(i-1)jk}) \\
\Delta_+^y o_{ijk} &= \frac{1}{h_y}(o_{i(j+1)k} - o_{ijk}) \\
\Delta_-^y o_{ijk} &= \frac{1}{h_y}(o_{ijk} - o_{i(j-1)k}) \\
\Delta_+^z o_{ijk} &= \frac{1}{h_z}(o_{ij(k+1)} - o_{ijk}) \\
\Delta_-^z o_{ijk} &= \frac{1}{h_z}(o_{ijk} - o_{ij(k-1)})
\end{aligned} \tag{76}$$

and the function $m(a, b)$ as:

$$m(a, b) = \frac{\text{sign}(a) + \text{sign}(b)}{2}\text{min}(|a|, |b|) \tag{77}$$

The function $sign(a)$ returns 1 for $x > 0$, -1 for $x < 0$ and 0 for $x = 0$. One advantage of $m(a, b)$ is that this function equals zero if a and b have the opposite sign. Thus, if $\Delta_+^{x,y,z} o_{ijk}$ and $\Delta_-^{x,y,z} o_{ijk}$ have not the same sign, we choose $\Delta_\pm^{x,y,z} o_{ijk} = 0$. We use the following boundary limits $\frac{\partial o}{\partial \vec{n}} = 0$, where $\vec{n}$ is normal to the boundary $\partial\Omega$ of Ω. For a discrete functional, this leads to:

$$\begin{array}{ccc}
o_{0jk} = o_{1jk}, & o_{i0k} = o_{i1k}, & o_{ij0} = o_{ij1} \\
o_{(Nx+1)jk} = o_{(Nx)jk}, & o_{i(Ny+1)k} = o_{i(Ny)k}, & o_{ij(Nz+1)} = o_{ij(Nz)}
\end{array} \tag{78}$$

Conjugate-Gradient Method

The parameters of the PSF are obtained by the Conjugate-Gradient method [66]. Analytically minimizing the log-likelihood w.r.t the parameters leads us to the following:

$$\begin{aligned}
\nabla_{\boldsymbol{\theta_j}} \mathcal{L}(\boldsymbol{\theta}_|) &= \frac{\partial}{\partial \boldsymbol{\theta}_j} \sum_{\mathbf{x} \in S} ((h * \hat{o})(\mathbf{x}) - i(\mathbf{x}) log[h * \hat{o}](\mathbf{x})) \\
\nabla_{\boldsymbol{\theta_j}} \mathcal{L}(\boldsymbol{\theta_j}) &= \sum_{\mathbf{x} \in S} ((\frac{\partial}{\partial \boldsymbol{\theta}_j} h * \hat{o}(\mathbf{x})) - i(\mathbf{x}) \frac{\partial}{\partial \boldsymbol{\theta}_j} log[h * \hat{o}](\mathbf{x})) \\
\nabla_{\boldsymbol{\theta_j}} \mathcal{L}(\boldsymbol{\theta_j}) &= \sum_{\mathbf{x} \in S} ((\frac{\partial}{\partial \boldsymbol{\theta}_j} h * \hat{o}(\mathbf{x})) - \frac{i(\mathbf{x})}{h * \hat{o}(\mathbf{x})} \frac{\partial}{\partial \boldsymbol{\theta}_j} h * \hat{o}(\mathbf{x})) \\
\nabla_{\boldsymbol{\theta_j}} \mathcal{L}(\boldsymbol{\theta_j}) &= \sum_{\mathbf{x} \in S} ((h_{\boldsymbol{\theta}_j} * \hat{o}(\mathbf{x})) - \frac{i(\mathbf{x})}{h * \hat{o}(\mathbf{x})} h_{\boldsymbol{\theta}_j} * \hat{o}(\mathbf{x}))
\end{aligned} \tag{79}$$

where,

$$h_{\boldsymbol{\theta}_j} = \frac{\partial}{\partial \boldsymbol{\theta}_j} h(\sigma_r, \sigma_z) \tag{80}$$

$$\frac{\partial}{\partial \boldsymbol{\theta}_j} h(\sigma_r, \sigma_z) = (-\frac{2}{\sigma_r} + \frac{r^2}{\sigma_r^3}) h(\sigma_r, \sigma_z); \boldsymbol{\theta}_j = \sigma_r \tag{81}$$

$$\frac{\partial}{\partial \boldsymbol{\theta}_j} h(\sigma_r, \sigma_z) = (-\frac{1}{\sigma_z} + \frac{z^2}{\sigma_z^3}) h(\sigma_r, \sigma_z); \boldsymbol{\theta}_j = \sigma_z \tag{82}$$

Hence,

$$\hat{\boldsymbol{\theta}_j}^{(k+1)} = \hat{\boldsymbol{\theta}_j}^{(k)} - \alpha^{(k)} \nabla_{\boldsymbol{\theta}_j} \mathcal{L}(\hat{\boldsymbol{\theta}_j}^{(k)}) \tag{83}$$

where $\alpha^{(k)}$ and $\nabla_{\boldsymbol{\theta}_j} \mathcal{L}(\hat{\boldsymbol{\theta}}^{(k)})$ are the step size and the search directions at iteration k.

References

[1] J. Hadamard. *Lectures on Cauchy's Problem in Linear Partial Differential Equations*. Yale University Press, New Haven, U.S.A., 1923.

[2] H. E. Keller. *Handbook of Biological Confocal Microscopy*, chapter Objective lenses for confocal microscopy, pages 111–126. Plenum Press, New York, 2nd edition, 1995.

[3] M. Born and E. Wolf. *Principles of Optics*. Cambridge Press, Cambridge, 1999.

[4] J. G. McNally, C. Preza, J.-Á. Conchello, and L. J. Thomas Jr. Artifacts in computational optical-sectioning microscopy. *J. Opt. Soc. Am. A*, 11:1056–1067, March 1994.

[5] J. B. Pawley. *Handbook of biological confocal microscopy*. Plenum Press, New York, 2nd edition, 1996.

[6] P. A. Stokseth. Properties of a Defocused Optical System. *J. Opt. Soc. Am. A*, 59:1314–1321, October 1969.

[7] S. Hell, G. Reiner, C. Cremer, and E.H.K. Stelzer. Aberrations in confocal fluorescence microscopy induced by mismatches in refractive index. *J. Microscopy*, 169:391–405, 1993.

[8] J. B. Pawley. *Three-Dimensional Confocal Microscopy: Volume Investigation of Biological Specimens*, chapter Sources of noise in three-dimensional microscopical data sets, pages 47–94. Academic Press, San Diego, California, 1994.

[9] B. Zhang. *Contributions to fluorescence microscopy in biological imaging: PSF modeling, image restoration, and super-resolution detection.* PhD thesis, École Nationale Supérieure des Télécommunications, Paris, 2007.

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

[11] G. B. Avinash. Data-driven, simultaneous blur and image restoration in 3-D fluorescence microscopy. *Journal of Microscopy*, 183(2):145–157, August 1996.

[12] L. Mandel. Sub-Poissonian photon statistics in resonance fluorescence. *Optics Letters*, 4:205–207, July 1979.

[13] N. Dey, L. Blanc-Féraud, C. Zimmer, P. Roux, Z. Kam, and J.C. Olivo-Marin. 3D Microscopy Deconvolution using Richardson-Lucy Algorithm with Total Variation Regularization. Research Report 5272, INRIA, France, July 2004.

[14] A. Buades, B. Coll, and J. M. Morel. A Review of Image Denoising Algorithms, with a New One. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005.

[15] C. Chaux, L. Blanc-Féraud, and J. Zerubia. Wavelet-based restoration methods: application to 3D confocal microscopy images. In *Proc. SPIE Conference on Wavelets*, San Diego, USA, August 2007.

[16] J. R. Monck, A. F. Oberhauser, T. J. Keating, and J. M. Fernandez. Thin-section ratiometric Ca2+ images obtained by optical sectioning of fura-2 loaded mast cells. *J. Cell Biol.*, 116(3):745–759, 1992.

[17] D. A. Agard. Optical sectioning microscopy: Cellular architecture in three dimensions. *Ann. Rev. Biophys. Bioeng.*, 13:191–219, 1984.

[18] A. Erhardt, G. Zinser, D. Komitowski, and J. Bille. Reconstructing 3-D light-microscopic images by digital image processing. *Appl. Opt.*, 24:194–200, 1985.

[19] T. Tommasi, A. Diaspro, and B. Bianco. 3-D reconstruction in optical microscopy by a frequency-domain approach. *Signal Process.*, 32(3):357–366, 1993.

[20] C. Preza, M. I. Miller, L. J. Thomas Jr., and J. G. McNally. Regularized linear method for reconstruction of three-dimensional microscopic objects from optical sections. *J. Opt. Soc. Am. A*, 9(2):219, 1992.

[21] A. N. Tikhonov and V. A. Arsenin. *Solution of Ill-posed Problems*. Winston & Sons, 1977.

[22] W.A. Carrington, K. E. Fogarty, and F. S. Fay. *3D Fluorescence Imaging of Single Cells Using Image Restoration*, chapter 3, pages 53–72. Wiley-Liss, 1990.

[23] T. J. Holmes. Maximum-likelihood image restoration adapted for noncoherent optical imaging. *J. Opt. Soc. Am. A*, 5:666–673, May 1988.

[24] S. Joshi and M. I. Miller. Maximum a posteriori estimation with Good's roughness for three-dimensional optical-sectioning microscopy. *J. Opt. Soc. Am. A*, 10(5):1078, 1993.

[25] G. M. P. Van Kempen, L. J. V. Vliet, P. J. Verveer, and H. T. M. V. D. Voort. A Quantitative Comparison of Image Restoration methods for confocal microscopy. *Journal of Microscopy*, 185:354–365, 1997.

[26] P. J. Verveer and T. M. Jovin. Efficient superresolution restoration algorithms using Maximum A Posteriori estimations with application to fluorescence microscopy. *J. Opt. Soc. Am. A*, 14(8):1696, August 1997.

[27] M.A.T. Figueiredo and R.D. Nowak. Wavelet-based image estimation: an empirical Bayes approach using Jeffrey's noninformative prior. *IEEE Trans. on Image Processing*, 10(9):1322–1331, Sep 2001.

[28] N. Dey, L. Blanc-Féraud, C. Zimmer, Z. Kam, P. Roux, J.C. Olivo-Marin, and J. Zerubia. Richardson-Lucy Algorithm with Total Variation Regularization for 3D Confocal Microscope Deconvolution. *Microscopy Research Technique*, 69:260–266, 2006.

[29] L. M. Bregman. The method of successive projection for finding a common point of convex sets (Theorems for determining common point of convex sets by method of successive projection). *Soviet Mathematics*, 6:688–692, 1965.

[30] M. Koshy, D. A. Agard, and J. W. Sedat. Solution of Toeplitz systems for the restoration of 3-D optical sectioning microscopy data. In Louis C. Smith, editor, *Bioimaging and Two-Dimensional Spectroscopy, Proceedings of the SPIE*, volume 1205, pages 64–71., 1990.

[31] P. J. Verveer, M. J. Gemkow, and T. M. Jovin. A comparison of image restoration approaches applied to three-dimensional confocal and wide-field fluorescence microscopy. *J. Microsc.*, 193:50–61, 1999.

[32] A.N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Sov. Math. Dokl.*, 4:1035–1038, 1963.

[33] K. Miller. Least Squares Methods for Ill-Posed Problems with a Prescribed Bound. *SIAM J. on Mathematical Analysis*, 1(1):52–74, 1970.

[34] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D.*, 60:259–268, 1992.

[35] P. Charbonnier, L. Blanc-Féraud, G. Aubert, and M. Barlaud. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. on Image Processing*, 6(2):298–311, Feb. 1997.

[36] M. Persson, D. Bone, and H. Elmqvist. Total variation norm for three-dimensional iterative reconstruction in limited view angle tomography. *Physics in Medicine and Biology*, 46(3):853–866, 2001.

[37] P. Charbonnier, L. Blanc-Féraud, and M. Barlaud. An adaptive reconstruction method involving discontinuities. *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 5:491–494, April 1993.

[38] G. Chenegros, L. M. Mugnier, F. Lacombe, and M. Glanc. 3-D deconvolution of adaptive-optics corrected retinal images. In J.-A. Conchello, C. J. Cogswell, and T. Wilson, editors, *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing*, volume 6090 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 144–151, March 2006.

[39] C. Vonesch and M. Unser. Fast Wavelet-Regularized Image Deconvolution. In *Proceedings of the International Symposium on Biomedical Imaging: From Nano to Macro*, pages 608–611, April 2007.

[40] J.-L. Starck and A. Bijaoui. Filtering and deconvolution by the wavelet transform. *Signal Process.*, 35(3):195–211, 1994.

[41] A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. Satellite image deconvolution using complex wavelet packets. Research Report 3955, INRIA, France, June 2000.

[42] R. M. Willett, I. Jermyn, R. D. Nowak, and J. Zerubia. Wavelet-Based Superresolution in Astronomy. In F. Ochsenbein, M. G. Allen, and D. Egret, editors, *Astronomical Data Analysis Software and Systems (ADASS) XIII*, volume 314 of *Astronomical Society of the Pacific Conference Series*, pages 107–116, July 2004.

[43] D. S. C. Biggs. Clearing up deconvolution. *Biophotonics Int.*, 11(2):32–36, 2004.

[44] J. W. Shaevitz and D. A. Fletcher. Enhanced three-dimensional deconvolution microscopy using a measured depth-varying point-spread function. *J. Opt. Soc. Am. A*, 24(9):2622–2627, 2007.

[45] S.F. Gibson and F. Lanni. Diffraction by a circular aperture as a model for three-dimensional optical microscopy. *J. Opt. Soc. Am. A*, A6:1357–1367, 1989.

[46] R. Hudson, J. N. Aarsvold, C.-T. Chen, J. Chen, P. Davies, T. Disz, I. Foster, M. Griem, M. K. Kwong, and B. Lin. Optical microscopy system for 3D dynamic imaging. In Carol J. Cogswell, Gordon S. Kino, and Tony Wilson, editors, *Proc. of SPIE*, volume 2655, pages 187–198, 1996.

[47] O. V. Michailovich and D. R. Adam. Deconvolution of medical images from microscopic to whole body images. In P.Campisi and K. K. Egiazarian, editors, *Blind Image Deconvolution: Theory and Applications*, chapter 5, pages 169–237. CRC Press, 2007.

[48] T. J. Holmes. Blind deconvolution of quantum-limited incoherent imagery: maximum-likelihood approach. *J. Opt. Soc. Am. A*, 9:1052–1061, July 1992.

[49] J. Markham and J.-A. Conchello. Parametric blind deconvolution: a robust method for the simultaneous estimation of image and blur. *J. Opt. Soc. Am. A*, 16(10):2377–2391, 1999.

[50] M. Jiang and G. Wang. Development of blind image deconvolution and its applications. *Journal of X-Ray Science and Technology*, 11:13–19, 2003.

[51] L. Bar, N. A. Sochen, and N. Kiryati. Variational Pairing of Image Segmentation and Blind Restoration. In Tomás Pajdla and Jiri Matas, editors, *Computer Vision-ECCV*, volume II, pages 166–177. Springer, May 2004.

[52] A. Santos and I. T. Young. Model-Based Resolution: Applying the Theory in Quantitative Microscopy. *Appl. Opt.*, 39(17):2948–2958, 2000.

[53] B. Zhang, J. Zerubia, and J. C. Olivo-Marin. Gaussian approximations of fluorescence microscope point-spread function models. *Appl. Opt.*, 46(10):1819–1829, 2007.

[54] T. J. Schulz. Multiframe blind deconvolution of astronomical images. *J. Opt. Soc. Am. A*, 10:1064–1073, May 1993.

[55] L.B. Lucy. An iterative technique for the rectification of observed distributions. *Astronomy*, 79:745–754, 1974.

[56] W.H. Richardson. Bayesian-Based Iterative Method of Image Restoration. *J. Opt. Soc. Am. A*, 62(1):55–59, January 1972.

[57] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society B*, 39(1):1–38, 1977.

[58] L. A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Trans. on Medical Imaging*, MI-1(2):113–122, Oct. 1982.

[59] C. J. R. Sheppard and C. J. Cogswell. Three-dimensional image formation in confocal microscopy. *Journal of microscopy*, 159(2):179–194, 1990.

[60] M. Jiang, G. Wang, M. W. Skinner, J. T. Rubinstein, and M. W. Vannier. Blind Deblurring of Spiral CT Images. *IEEE Trans. Med. Imaging*, 22(7):837–845, 2003.

[61] I. Csiszar. Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems. *The Annals of Statistics*, 19(4):2032–2066, 1991.

[62] P. Pankajakshan, B. Zhang, L. Blanc-Féraud, Z. Kam, J.-C. Olivo-Marin, and J. Zerubia. Parametric Blind Deconvolution for Confocal Laser Scanning Microscopy. In *Proc. 29th International Conference of IEEE EMBS (EMBC-07)*, Lyon, France, August 2007.

[63] 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. Technical report, INRIA Sophia-Antipolis, Sophia-Antipolis, France, March 2008.

[64] A. Jalobeanu, L. Blanc-Féraud, and J. Zerubia. Hyperparameter estimation for satellite image restoration using a MCMC Maximum Likelihood method. *Pattern Recognition*, 35(2):341–352, 2002.

[65] D. Zhu, M. Razaz, and R. Lee. Adaptive penalty likelihood for reconstruction of multidimensional confocal microscopy images. *Computerized Medical Imaging and Graphics*, 29(5):319–331, July 2005.

[66] K. E. Atkinson. *An introduction to Numerical Analysis*. John Wiley and Sons, 2nd edition, 1989.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399