3D compact timelapse microscope through intensity diffraction tomography coupled with deep neural networks.

William Pierre

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ GRENOBLE ALPES

École doctorale : MSTII - Mathématiques, Sciences et technologies de l'information, Informatique
Spécialité : Mathématiques Appliquées
Unité de recherche : Institut National de Recherche en Informatique et en Automatique

Vidéo-microscope 3D et compact, utilisant de la tomographie de diffraction optique sans mesure de phase couplé à des réseaux de neurones profonds.

3D compact timelapse microscope through intensity diffraction tomography coupled with deep neural networks.

Présentée par :

William PIERRE

Direction de thèse :

Sergei GRUDININ	Directeur de thèse
Lionel HERVE Ingénieur-Chercheur (CR), CEA Grenoble	Co-directeur de thèse

Rapporteurs :

OLIVIER HAEBERLé
PROFESSEUR DES UNIVERSITES, Université Haute-Alsace
Bruno MONTCEL
MAITRE DE CONFERENCE, Université Lyon 1

Thèse soutenue publiquement le **29 novembre 2022**, devant le jury composé de :

Michel DOJAT DIRECTEUR DE RECHERCHE, Inserm	Président
Corinne FOURNIER MAITRE DE CONFERENCE, Université Saint-Étienne	Examinatrice
Anne SENTENAC DIRECTEUR DE RECHERCHE, CNRS	Examinatrice
OLIVIER HAEBERLé PROFESSEUR DES UNIVERSITES, Université Haute-Alsace	Rapporteur
Bruno MONTCEL MAITRE DE CONFERENCE, Université Lyon 1	Rapporteur
Ronald PHLYPO MAITRE DE CONFERENCE, Grenoble INP	Examinateur

Invités :

Lionel HERVE
INGENIEUR HDR, CEA
Sergei GRUDININ
CHARGE DE RECHERCHE, Inria Grenoble

DEDICATION

I want to express my appreciation to all the jury member for taking time to evaluate my work. I hope that this work will keep you enthralled in the same way it motivated me for three years. I am looking forward to hear your feedbacks during my PhD defense. Considering most of my gratitude goes to french people, I'll switch in french just after greeting my PhD supervisor. Despite the pandemic, Sergei was always here for any inquiry, he gave me numerous ideas, some of which had a deep impact on the thesis.

Tout d'abord merci Lionel, mon co-encadrant, pour la remise à niveau plus que nécessaire en optique ondulatoire et en optimisation, pour les discussions souvent variées et parfois farfelues, mais toujours enrichissantes !

Ensuite merci Sophie pour l'accueil au LSIV, pour tes conseils et ta souplesse sur certains protocoles complexes en ces temps troublés qui m'ont permis de continuer à travailler dans de bonnes conditions. Je tiens aussi à remercier toutes les personnes sans qui cette thèse n'aurait été possible :

- Cédric, pour avoir toujours mis en avant le fait que j'étais bien là pour faire de la recherche et non pour faire avancer un projet interne, merci de m'avoir transmis ta vision de la recherche, intègre.

- Ondrej pour m'avoir considérablement aidé en optique, à comprendre comment fonctionnait un objectif et comment faire des montages propres ! Le début de cette thèse aurait été bien plus laborieux sans toi.

- Magali pour m'avoir accueilli à bras ouvert à chacun de mes passages au CHU, pour m'avoir expliqué et réexpliqué la différence entre un zygote, un ovocyte, un blastocyste, un morula (promis je sais maintenant) Merci aussi pour tous les échantillons que tu m'as fournis sur ton temps libre et même sur tes dimanche (ces protocoles biologiques qui doivent être exécutés à heure fixe sont terribles)

- Chiara pour tout le soutien que tu m'as apporté pendant des heures en m'aidant

à préparer mes oraux et mes écrits, en rendant mon anglais moins “colloquial”. Si ce manuscrit a cette qualité-là, c’est en grande partie grâce à toi. Merci aussi pour toutes les discussions et propositions d’améliorations qui ont beaucoup apporté au prototype, continue à la chouchouter, en tout cas le laboratoire a beaucoup de chance de t’avoir.

- Laurence pour m’avoir guidé dans toutes les démarches CEA, aussi complexe soit-elle, merci pour ta patience dans les explications et ré-explications !

Je pense qu’il est aussi important de remercier les personnes sans qui cette thèse n’aurait pas été complètement impossible, mais surement beaucoup moins agréable :

- Caroline, la meilleure voisine de bureau imaginable, surtout quand il s’agit d’essayer de traduire l’un des 7043 acronyme du CEA, c’est quoi déjà une DAE et un PEPR ?

- Tous les jeunes (et moins jeunes) du labo qui mangent au H3 (le H1 c’est trop loin) le midi, merci pour les discussions légères et les bons moments !

De manière plus personnelle je voudrais aussi remercier ma famille, plus particulièrement mes parents pour le soutien sans faille qu’ils m’apportent et ce quel que soit mes projets ! Merci papa, merci maman, pour tout et surtout pour les framboises du jardin ! Merci les bros’ pour m’avoir, je pense, donné le gout des sciences et l’envoie d’aller faire ingénieur sans trop savoir ce que ce mot voulait dire.

Enfin merci à toi Anne, je t’aime, de tout mon cœur.

ABSTRACT

We aim at retrieving 3D refractive optical index (RI) of large samples (around 200 µm) from Intensity Diffraction Tomography (IDT) where the dataset is obtained by recording diffraction image with a variety of illumination angles. IDT targets biological imaging in a label-free manner using the optical variation within the sample and multiple tilted imaging to reconstruct the 3D map of RIs. However, standard IDT techniques reveal several drawbacks in terms of limited field of view and feasibility of imaging living samples in time-lapse conditions. We focused on time-lapse imaging of large sample (>200µm) without the need of large NA objective or immersion oil.

The challenge created by the absence of the phase information (intensity only measurements) as well as the limited illumination angle (low NA due to low magnification) has been solved using a Beam Propagation Method (BPM) embedded inside a deep leaning framework. The layers are encoding the 3D optical representation of the sample. Besides, we included in the forward model the effect of the spherical aberration introduced by the optical interfaces, which gave a strong impact on measurements under oblique illumination in terms of 3D spatial resolution. Using this framework, we achieved 3D reconstructions of mouse embryos (>100µm) in time-lapse conditions over 7 days, as well as liver organoid. We could observe the intrinsic embryonic development from single cell (low-scattering sample) to the blastocyst level (highly scattering sample) as well as the rotational movement and growth rate of liver organoid. Such time-lapse yields quantitative information on the development and viability of biological sample in view of the sub-cellular imaging capacities. Our technology opens up novel opportunities for 3D live cell imaging of whole organoids in time-lapse.

However, the use of a wide field of view inherently limits the maximum illumination angle possible. This directly affects the 3D reconstruction as it generates an axial elongation of each object or inner substructure, leading to a poor axial resolution (10 times greater than radial resolution) Motivated by the great results of Convolutional Neural Network

in correcting reconstruction problem, we decided to apply deep learning methods to our elongation problem. The idea of our process is to combine two learning methods each proven to solve a different axial reconstruction issue. One network recovers the inner structures; the other retrieves the global structure of the object. To ensure the continuity of the CNN output along each direction the training is performed on 3D random synthetic volumes. We will assess the quality of this learning method by checking its efficiency on mouse embryo and prostatic organoid as well as many various synthetic data.

RESUME

Le but de cette thèse est de reconstruire en 3D l'indice de réfraction de gros échantillons biologiques (> 200 µm) grâce à de la tomographie de diffraction à partir d'image d'intensité (IDT). L'information 3D est encodée dans une multitude d'images de diffraction à des angles d'éclairage différents.

La tomographie de diffraction n'a pas besoin de fluorescence pour imager des échantillons biologiques, elle se contente d'exploiter les changements de propriétés optiques de l'échantillon pour en reconstruire une carte 3D de l'indice de réfraction. Comme les techniques standards ont des champs de vue limités ainsi que des systèmes complexes souvent non compatibles avec les incubateurs, on s'est principalement concentré sur la création d'un système compact pouvant imager de gros échantillons (> 200 µm) sans huile de couplage ni objectifs à forte ouverture numérique.

La complexité de la reconstruction 3D causée par une faible couverture angulaire (elle même causée par une faible ouverture numérique d'objectif à faible grossissement), un manque d'information (Image d'intensité uniquement, sans la phase) est grande. Afin d'y parvenir, un modèle multicouche prenant en compte les diffractions multiples (le BPM) a été implémenté dans une infrastructure logicielle d'apprentissage profond. Les aberrations sphériques générées par des épaisseurs de verres non standards sont aussi prises en compte dans le modèle direct. Cela nous a permis d'imager toutes les heures pendant sept jours le développement d'un embryon de souris du stade unicellulaire au stade blastocyste. Ces capacités d'imagerie ont aussi été validées sur des organoïdes de foie. Ces timelapses apportent des informations quantitatives sur le développement et la viabilité des échantillons biologiques. Cette technologie apporte de nouvelles possibilités pour l'imagerie 3D en incubateur.

Cependant, l'utilisation d'un grand champ de vue limite la couverture angulaire exploitable. Cela affecte grandement les reconstructions 3D, qui se retrouvent étirées axialement, dégradant cette résolution. Motivé par de récentes utilisations de réseaux de

neurones convolutifs dans le domaine de la correction d'artefact de reconstructions, nous avons décidé de résoudre ce problème à l'aide d'une approche moderne par apprentissage profond. L'idée principale est de combiner deux réseaux, un qui retrouve la forme originale des structures intracellulaires, et un qui estime la forme extérieure de l'objet. Les réseaux sont entraînés sur données simulées uniquement. Ces données sont produites avec des objets 3D aléatoires mimant de possibles formes d'objets multicellulaires.

La généralisation de cet ensemble de réseaux est testée avec succès sur l'embryon à différents stade de développement, mais peine à se généraliser aux organoïdes de foie.

Contents

List of Figures

List of Symbols and Abbreviations

3D	Three dimensional
RI	Refractive index
ODT ...	Optical Diffraction Tomography
IDT	Intensity Diffraction Tomography
QPI	Quantitative phase imaging
PALM .	PhotoActivated Localization Microscopy
STORM	Stochastic Optical Reconstruction Microscopy
FPALM	Fluorescence PhotoActivation Localization Microscopy
CNRS .	Centre National de la Recherche Scientifique
LS	Lippmann Schwinger
BPM ...	Beam propagation method
MBS ...	Multiple Born Scattering
PDE ...	Partial Differential Equation
GD	Gradient Descent
TV	Total Variation
DL	Deep Learning
CPU ...	Central Processing Unit
GPU ...	Graphical Processing Unit
ADMM	Alternating Direction Method of Multipliers
MRI ...	Magnetic Resonance Imaging

Chapter 1

Introduction

We live in a three dimensional world that is constantly evolving, changing and moving. We started to observe and understand the world around us with our own eyes before recently creating sensor to extend our own limited capabilities. Nowadays our synthetic perception range goes from huge galaxy millions of kilometers away to individual atom. This inaccessible information for our primitive senses was revealed thanks to a tremendous amount of work in physics, engineering and computational power. Such progress have radically changed our understanding of the world.

Microscopes are a part of the story and are deeply link to the stereotypes of scientist, if we type science in google image, 25% of the images include a microscope on the first page. The idea of seeing small object that are around us at every moment but that cannot be seen with our bare eyes might explain this collective psyche. Looking through a microscope is a way to ask a fundamental question about the life that surround us. The microscope characteristics such as field of view or resolution mark the boundary of what can be questioned and more importantly how precise will our answer be.

Over the past decade, microscopes became more sophisticated, more complex and more demanding on the studied sample. In this thesis we will focus on gentle imaging technique that are soft with the sample because we want to study the development of

biological sample trough time without changing habits of biologist or perturbing their experiments. Organoids or cell cultures are grown in incubator that are often not bigger than a small fridge forcing us to use simple and compact hardware. This was made possible only because of the progress in Computed Imaging where the final images are extracted from the measurements using algorithms that rely on a significant amount of computing. The microscope encodes the desired information, and the computational power is required to decode the information.

Every Computed Imaging system is a set of three intrinsic parts that must be designed as a whole. It is always composed of:

- An imaging system that can encodes the information
- A physical model of the imaging system
- A reconstruction algorithm to decode the desired information

This introduction is here to give a context to this thesis, starting with a brief history of standard microscopy, before diving into 3D imaging of biological sample with fluorescent markers. Finally, I will review the standard methods used in 3D phase imaging system that relies on computed imaging system.

1.1 FROM STANDARD MICROSCOPY TO COMPUTATIONAL IMAGING

Before the invention of light microscopy, lenses had been produced for many centuries, but their entire scientific potential had remained hidden. It all started in 1590 when Hans and Zacharias put lenses in a tube to magnify object [1]. This invention remained unnamed and we had to wait 1609 for Galileo to design its famous compound microscope and 1625 for Giovanni Faber to formulate the word describing Galileo invention: *microscope*.

Microscopes where then used for decades to unveil new insight in microbiology with the invention of the word *cell* by Anton van Leeuwenhoek. He was known for polishing and grinding his own lenses, greatly improving microscope quality at that time, enabling

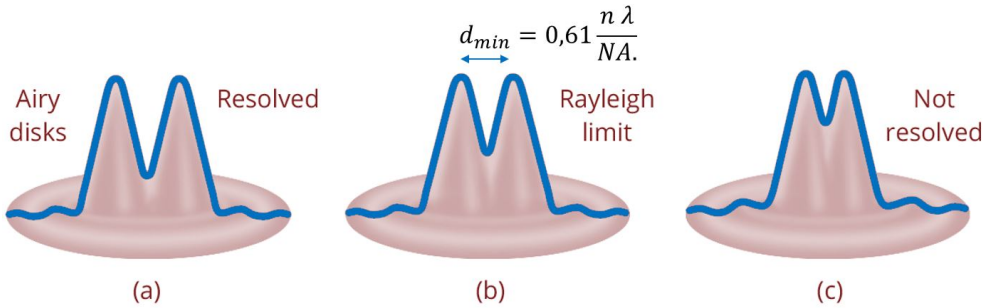


Figure 1.1: The Raleigh limit of resolution. NA is the numerical aperture of the objective, λ the wavelength of the light source and η the refractive index between the object and the objective.

him to see the many microscopic animals [49] and bacteria surrounding us. Even though microscopic technique improved with the creation of Achromatic Lens in the 18th century, their true theoretical modelling and technical limitation were introduced in the nineteenth century with the famous diffraction limit theory.

The resolution of conventional light microscopes has an inherent limit; by working in the "far-field", *i.e.* many wavelength away from the sample, the light diffraction becomes significant. Using a microscope makes the feature of interest in the sample to scatter light, this scattering creates blurry patches called Airy disks in Fig.1.1. The generally accepted criterion for the minimum resolvable detail before two objects blur into each other is the Raleigh criterion [36].

Nowadays CMOS sensor are still limited by the Raleigh criterion but they are more and more efficient and capture directly the electromagnetic signal with a low noise to signal ratio compare to our eyes. A signal that can be analyzed, filtered, or modified computationally to produce better image. Computational part emerged in every part of the microscope, whether it is to control the auto-focus [77] [74] or the fluorescence wavelength. Such *computationally-assisted microscopes* were optimised to be used through a computer and not by a human operator. Even though it was not meant to be used by human, the output of such microscope were high quality image directly understandable by a human.

On the contrary, *computational microscopes* need on top of the microscope a physical

model of the setup as well as a reconstruction algorithm to produce human readable data output. A simple example of a *computational microscopes* is the simplest modern microscope possible the lens free microscope. The sample is placed on top of a CMOS sensors bellow a LED without any lenses, the raw image is the scattering pattern created by the sample as shown in Fig.1.2. Using a free space propagation model it is possible to compute back what was the object that produces such pattern.

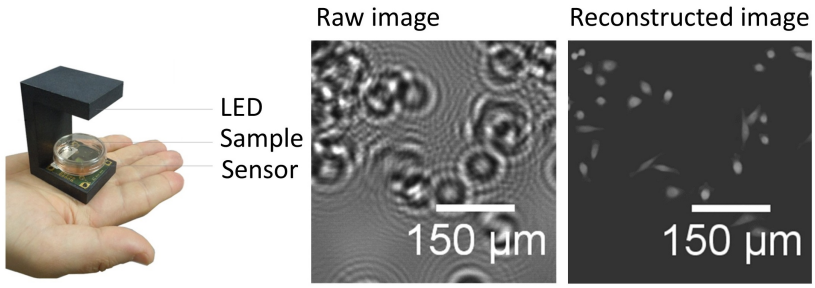


Figure 1.2: Example of a computed microscope The lens free microscope with its simple setup.

1.2 BIOLOGICAL MICROSCOPY THANKS TO FLUORESCENCE IMAGING

Microscopes first intensive usage was for biological research with the publication of *Micrographia* in 1665 [49]. It was then used in many other field such as material characterization [52] or monitoring of semiconductor chip manufacturing [113], but the main usage remained biological imaging.

One of the key discovery that put microscopy as reference technique for biologist is fluorescence labeling [102]. Fluorescent labels are nowadays the gold standard to label cell unit depending on their intrinsic chemistry to unveil the interaction between various cell parts inside a more complex organism. Fluorescence labeling binds fluorescent dyes to functional groups contained in biological molecules so that they can emit light when they are properly excited. The re-emitted light has a larger wavelength and this facilitates the imaging of the desired bio-molecule. The emergence of various fluorophores has increased the number of possibilities for the sensitive detection of bio-molecules and mul-

tiple fluorophores may be used simultaneously thus enabling biologist to analyse complex interactions. It produces stunning results but it has also drawbacks. The fluorescence degrades over time: photobleaching occurs, *i.e.*, fluorophores emit less and less light when they are repetitively illuminated. The process to label the desired object is time consuming. Finally, fluorophores only bind to a small amount of the desired cell substructures.

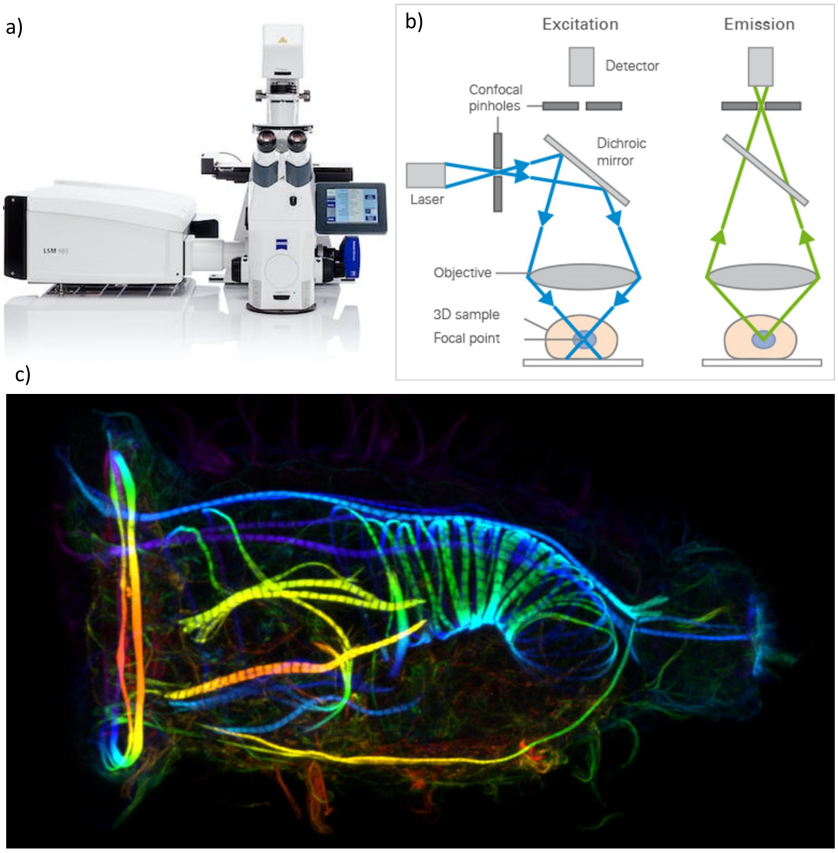


Figure 1.3: Illustration of the confocal microscopy. a) Commercial microscope ZEISS LSM 980 b) Excitation and emission light pathways in a basic confocal microscope configuration. c) Muscles of a Cyphonautes (aquatic invertebrate animal larva). Courtesy of Dr. Bruno C. Vellutini for Nikon Small World 2018

Despite these drawbacks, its high signal to noise ratio and its single molecule high sensitivity maintain fluorescence imaging as a standard. The three main techniques are the confocal microscopy, the light-sheet microscopy and the multi-photon microscopy. All of these techniques are intricately 3D and since most biological tissues and cells are inherently

spatial (three-dimensional objects), it is particularly suited to biological imaging.

1.2.1 3D fluorescence imaging technique

The first used 3D microscopy imaging technique is confocal imaging, presented in Fig.1.3. Its goal is to get rid off light that comes from out of focus plane. It is achieved with a diaphragm in the conjugate focal plane [115] [2] [23](explaining the name "confocal"). The volume to analyse is sensored one point at a time leading to time consuming acquisition for wide 3D volume at high resolution. They remain widely used nowadays because of their ease of use and manufacturing.

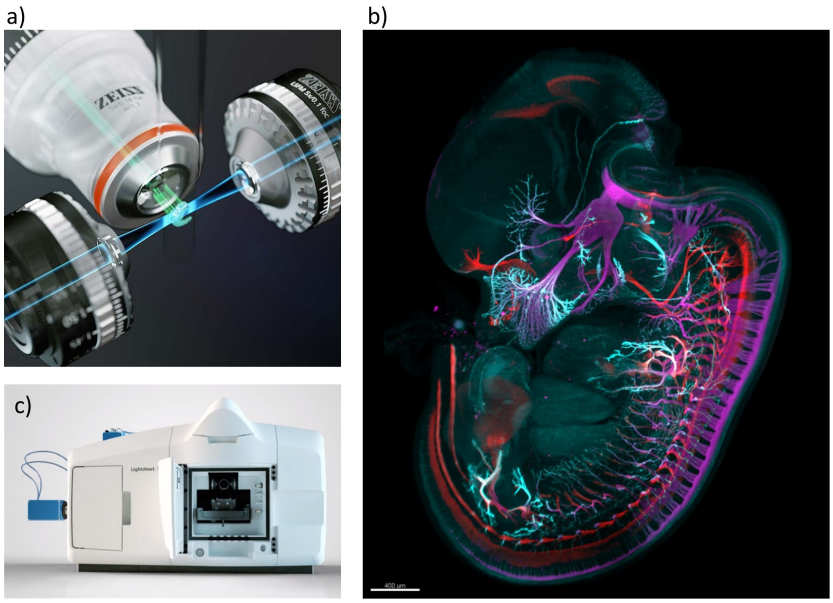


Figure 1.4: Illustration of the light sheet microscopy. a) Excitation light shaped in a thin plane is in Blue and the emitted fluorescence light is in green. b) Commercial microscope, ZEISS lightsheet 7 . c) Mouse embryo (day 12.5) stained for motor (red) and sensory (magenta) nerves and nerve endings (cyan). Courtesy of Dr. Gist F. Croft Lauren Pietila, Dr. Ali H. Brivanlou for Nikon Small World 2018

Multiple improvement have been brought to confocal imaging in order to reduce the acquisition time or the photo bleaching of fluorophores due to the high amount of incident light required. The idea emerged to illuminate only a plane of the sample with sheet shaped laser beam. In light sheet microscopy [134] [4] [124] the beam is shaped with a

cylindrical lens into a thin plane orthogonal to imaging part as shown in Fig.1.4. The 3D volume is created by stacking the imaged plane. The main advantages are to decrease the imaging time as well as increasing the time of imaging for a sample. These setups are more complex to build but the first commercial products arrived in the 2010s.

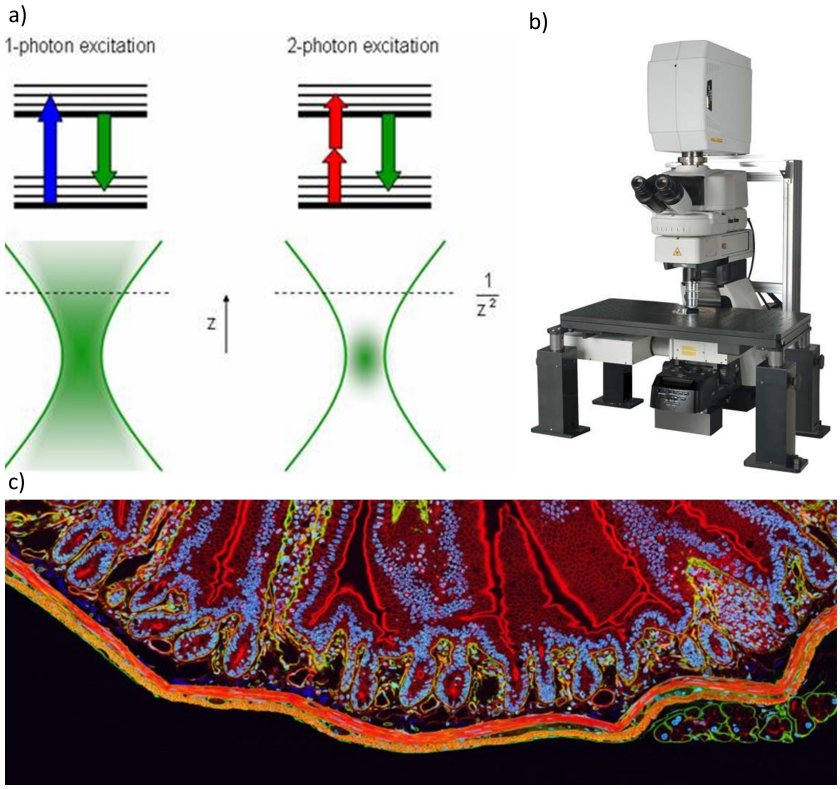


Figure 1.5: Illustration of a two photon microscopy. a) Schematics of a one and two photon excitation strategy. b) Commercial microscope. c) Quantum dot fluorescence image of mouse small intestine. Courtesy of Thomas J. Deerinck for photomicrography competition 2005

Other way of improving confocal imaging also emerged with two-photon [27] [28] [146] and multi photon imaging [42] [29]. The idea behind this technique illustrated in Fig.1.5 is that if a fluorophore can absorb a photon with wavelength λ_1, it can also absorb 2 photons simultaneously that have a wavelength λ_2 as long as $\lambda_1 = 2\lambda_2$ and they arrive at the same place, at the same time. This constraint is mostly respected in the focal plane of the illuminated objective where the photon density is the highest. The acquisition time

remains unchanged, but the contrast is largely increased since out of focus fluorescence is merely excited. Another good point in that longer wavelength are used, enabling a deeper penetration of the illumination light. Finally, the sensitivity of the microscope is increased since aperture in the conjugate focal plan is no longer required. Finally, less light is used, so photo-bleaching is reduced.

All these fluorescence imaging technique remained limited by *the diffraction barrier* until 2006, the year of a general emergence of super-resolution microscopy for fluorescence imaging. PALM [9], STORM [109], FPALM [46] enable the computation of images way beyond the diffraction limit by collecting a large amount of images, each having only some isolated activated fluorophores as shown in Fig. 1.6.

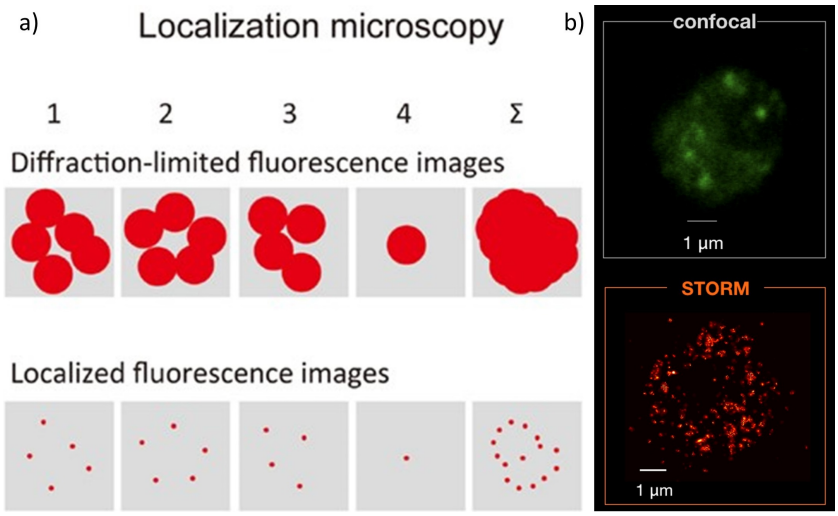


Figure 1.6: Photo-Activable Localization Microscopy illustration a) Scheme of the PALM technique b) Confocal and STORM images of nuclear bodies in Drosophila cells (unpublished). Courtesy of CBS from CNRS

Biological fluorescence microscopy became a standard for the study of, often chemically cleared, cell cultures, organoids or embryos. Even though fluorescence microscopy is widely used, it has inherent drawbacks.

- Natural auto-florescence is rare and most of the time the samples have to be labeled or worst, genetically modified. It is then never possible to take images from the

sample in its natural state.

- The dyes injection in the sample must be carefully executed with strict biological protocols that are time- consuming.
- Fluorophores are not neutral in the biochemistry of the sample, they may induce toxicity for two reasons. Either because they release toxic molecules for the sample, or because their activation damages the sample tissue or molecules.
- Only some parts of a cell of organoid can be functionally activated with florescent dyes, inherently limiting fluorescence imaging to theses parts as the other one a not being observable.
- Fluorescence gives information about shape and density, but its is not a quantitative imaging technique, unless special specific criteria are met [91].
- The number of fluorophores available is low and they respond to high energy photons (ultraviolet) that cause phototoxicity.
- The performance of a fluorophore degrades with time; it can be problematic when studying a sample for long periods of time.

To overcome these drawbacks, other imaging techniques have emerged, like the quantitative phase imaging (QPI).

1.3 QUANTITATIVE PHASE IMAGING TOWARD INTENSITY DIFFRACTION TOMOGRAPHY

Quantitative phase imaging relies on the delays an illuminating field encounters as it passes through the sample to reconstruct the structural information of the imaged sample.

The incident light encounters longer travel lengths in dense materials and the sample's non-uniform 3D distribution of its refractive index relative to its surrounding medium will create temporal delays. Those delays create scattering pattern in the far field.

Depending on the intensity of the refractive index gap, the generated scattering can be light or very strong to the point it can perturb significantly the illumination. Directly measuring the phase of a incoming visible light field is not possible, that is why quantitative phase imaging focuses on the delay introduced by the sample, measuring only the phase shift ϕ compared to a reference field.

$$\phi = \frac{2\pi}{\lambda}\delta n D$$

Here D stands for the thickness of the sample, δn if the refractive index gap with the refractive index around the sample, λ is the wavelength of the incoming light.

Quantitative phase imaging is complex for two reasons; first the phase information that is our value of interest is not accessible with common image sensor that only record the intensity of light and not its phase. This can be solved with interferometric techniques for example [62]. The second issue is the difficulty to separate the refractive index gap δn from the sample thickness D when you have managed to record the phase-shift ϕ.

To overcome these complications QPI uses custom setup that encodes the phase information inside the intensity image and that includes also 3D information about the object to enable a 3D reconstruction of the refractive index gap [56].

These 3D QPI methods have been welcomed by the biology community as refractive index imaging relies exclusively on the 3D shape of the imaged sample without having to use contrast agents. This advantage over standard methods like fluorescence imaging is important in the sense that the sample can be studied intact and not tampered by any kind of toxic fluorescent dyes. This is crucial to truly understand biological life in its "natural" condition.

The most mature system for 3D QPI is the Optical Diffraction Tomography (ODT).

It uses scanned illumination angle coupled with an interferometric detection system to directly capture phase information in the intensity measurements [139] [89] [126]

In the recent year, ODT's hardware and reconstruction algorithms were improved, widening its range of possibility in a compact and robust hardware that is commercialized. The recent developments focused on the use of digital micromirror devices [117], annular illumination [21] and reconstruction improvements through iterative [80] and learning-based reconstruction methods [58] and automatic regularisation parameter tuning [25]. Most of this advancement made it possible for this technique to reach commercial success [3] with application in Cell biology [116] [111], Microbiology [57] [63] etc. ...

Though ODT is the 3D QPI technique the most widely used, it has some drawbacks that limit its utility in biological research. The first one comes from its use of laser-based illumination for scanning the sample. Lasers have high spatial coherence and long temporal coherence lengths that generates high quality interference fringes that enable to retrieve the sample's phase with high sensitivity. From this laser illumination comes coherent noise, phase instabilities and complex system alignment protocol. The coherent noise comes from unwanted interference that take the form of speckle. Moreover two-arm interferometer design can introduce phase offsets and inaccurate phase measurements if there are any route length anomalies. This severe geometric constraint is only respected with very stable mechanical construction. Otherwise an extensive calibration procedure is needed to produce accurate 3D QPI, these constraints may stop biological labs to build custom ODT installations. Commercial setup of ODT provides pre-calibrated, well-designed setups ready for 3D QPI, are still costly.

To overcome these limitations and make 3D QPI more accessible, a great amount of work was done with standard transmission microscope and uncomplicated hardware modification. Intensity-only 3D QPI techniques, also known as Intensity diffraction Tomography (IDT) get rid of the phase information from the measurement in exchange of a simplified and more robust setup, like in Fig.1.7. The phase information is retrieved from a variety of possible phase encoding strategies like sample scanning or rotating [39] [55]

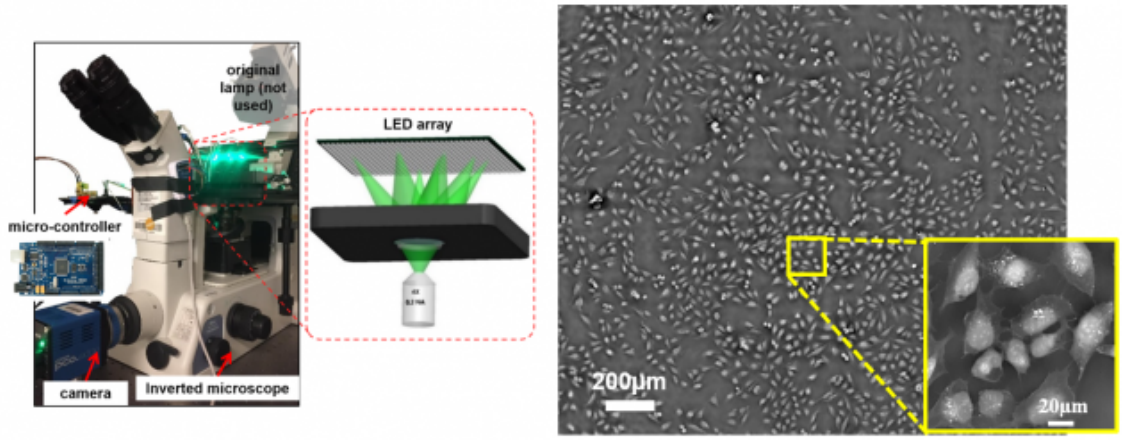


Figure 1.7: Modified inverted microscope for quantitative phase imaging from intensity image only. Courtesy of Tianlab

[121], custom illumination [75] [13], pupil engineering [89] [138] etc. If the sample produces strong multiple scattering, iterative reconstruction algorithm with multiple scattering physical model [128] [15] [78] [58] [96] were developed. The computational times increases strongly but with the emergence of GPU and efficient implementation of these multiple scattering model reconstruction can happen in quasi real time for time-lapse imaging.

1.4 THESIS OUTLINE

This thesis aims at providing a gentle microscope with a wide field of view ($> 100 \times 100 \times 100\ \mu m^3$) to image wide and thick biological samples. Fluorescence imaging is not an option here because we want our microscope to be non-intrusive and to keep the sample integrity. Quantitative phase imaging system are aligned with our vision of label free microscopy, but in order to enhance the simplicity and usability of our imaging system for biological study we want to keep the hardware as simple (and cheap) as possible to maximise its range of usage possibility. That is why interferometric system were not used as well as laser-based illumination.

Considering these constraints on the absence of staining agent, field of view, capacity, resolution and imaging frequency we decided to design a LED array microscope, where

the 3D distribution of the refractive index of the object is encoded in multiple intensity only images for various illumination angles. Chapter 2 discusses the scattering of light with various models that goes from the standard and simple born approximation toward an iterative solving of the Helmholtz equation. A multi-slice model that accounts for multiple scattering was chosen due to its higher performance for the desired discretization size.

Chapter 3 introduces some optimisation algorithms to recover the 3D refractive optical index volume of an object from multiple angle intensity-only-measurement. The algorithms are formulated as inverse problems, and solutions are obtained through an optimization performed by gradient descent based solvers. We model the physics of light scattering by using a deep learning framework, enabling automatic gradient propagation and GPU acceleration. The reconstruction algorithm is validated on simulated and biological data. Finally, the time-lapse ability of the whole system is also validated with a week-long 3D video of an mouse embryo developing.

Chapter 4 introduces Deep Learning with a short history and its key features that make it so appealing in computer vision. First, a failure case is presented on the training of a 2D Unet on random stack of random image. The complete reconstruction loop is applied to the generated volume to create pairs of perfect and artefact full object. The neural network then tries to go from the corrupted images to the real images.

A second attempt on synthetic 3D continuous volumes was successfully applied to the complete time-lapse of mouse embryo.

Chapter 2

Monochromatic light scattering physics and GPU implementation

The goal of this chapter is to show how to model the diffraction of plane wave by the variation of the complex refractive index of a sample.

Multiple models have been introduced to describe the way light interacts with changing refractive index. This chapter will start with the Helmholtz equation that is the problem we are trying to solve and a brief reminders on plane wave and Fourier transform that is fundamental before doing physical optics.

Then a reference model will be introduced, The LS model that solves interactively the Helmotz equation. This model is complex to evaluate and takes a lot of computing power but it is very accurate even with strongly scattering object.

Finally the usual models used in optical tomography will all be compared to this reference model to choose the most precise model for our discretization size.

2.1 FROM MAXWELL TO HELMHOLTZ

Lets consider a space Ω centered on O with direct orthonormal X, Y, Z axes. Any point of this space will be noted $\vec{r} = (x, y, z)$. If $\epsilon(\vec{r})$ and $\mu(\vec{r})$ are respectively the permittivity

and the permeability of Ω, macroscopic formulation of Maxwell equations in the absence of free charges are:

$$\vec{\nabla}.\vec{E} = 0 \tag{2.1}$$

$$\vec{\nabla}.\vec{H} = 0 \tag{2.2}$$

$$\vec{\nabla} \times \vec{E} + \mu \frac{\partial \vec{H}}{\partial t} = 0 \tag{2.3}$$

$$\vec{\nabla} \times \vec{H} - \epsilon \frac{\partial \vec{E}}{\partial t} = 0 \tag{2.4}$$

where $\vec{E}$ is the electric field, and $\vec{H}$ is the magnetic field. More precisely, the rectilinear components of $\vec{E}$ on X, Y, Z are respectively (E_x, E_y, E_z). Same notation will be used with $\vec{H} : (H_x, H_y, H_z)$. Both $\vec{E}$ and $\vec{H}$ depends of the time t and the position $\vec{r}$.

For biological samples, media are generally dielectric, linear, isotropic so we will limit our study to media with such properties. The linearity of the medium is satisfied if the medium has all the usual linearity properties [40]. The isotropic property of the medium is satisfied if ϵ and μ are independent of the polarization direction ($\vec{E}$ and $\vec{H}$ direction). The nondispersive property is satisfied if ϵ is not function of the wavelength in a wavelength region in which the wave propagates.

By applying operator $\nabla\times$ on both side of Eq.2.2, by using the vector identity ($\vec{\nabla} \times \vec{\nabla} \times \vec{A} = \vec{\nabla}(\vec{\nabla}.\vec{A}) - \vec{\nabla}^2\vec{A}$) and Eq.2.4, we get:

$$\vec{\nabla} \times (\vec{\nabla} \times \vec{E}) + \mu \frac{\partial \vec{\nabla} \times \vec{H}}{\partial t} = \vec{\nabla}(\vec{\nabla}.\vec{E}) - \vec{\nabla}^2\vec{E} + \mu\epsilon \frac{\partial^2 \vec{E}}{\partial t^2} = 0 \tag{2.5}$$

where c is the vacuum propagation velocity defined by the relation $\epsilon_0\mu_0c^2 = 1$. Eq.2.5 can

be expressed with only c and the refractive index $\eta = \sqrt{\epsilon/\epsilon_0}$ as:

$$\vec{\nabla}^2\vec{E} - \frac{\eta^2}{c^2}\frac{\partial^2\vec{E}}{\partial t^2} = 0 \tag{2.6}$$

If we apply the same calculus to $\vec{H}$ we find the same exact equation:

$$\vec{\nabla}^2\vec{H} - \frac{\eta^2}{c^2}\frac{\partial^2\vec{H}}{\partial t^2} = 0 \tag{2.7}$$

We can further notice that each component of $\vec{H} : (H_x, H_y, H_z)$ and $\vec{E} : (E_x, E_y, E_z)$ obeys to the same scalar equation, for example with H_x:

$$\Delta H_x - \frac{\eta^2}{c^2}\frac{\partial^2 H_x}{\partial t^2} = 0 \tag{2.8}$$

Therefore, we summarize the behaviour of the electromagnetic field by a single scalar wave equation:

$$\Delta U(\vec{r}, t) - \frac{\eta(\vec{r})^2}{c^2}\frac{\partial^2 U(\vec{r}, t)}{\partial t^2} = 0 \tag{2.9}$$

where $U(\vec{r}, t)$ can be any of the electromagnetic field components.

This approximation is known as the scalar theory and it is widely used in optics, it relies on the hypothesis that the spatial gradient of ϵ is small. As stated in the 3^{rd} chapter of *Introduction to Fourier Optics* [41], this approximation remains valid as long as the studied structures are small compared to the wavelength.

As we aim at analyzing signal from diffraction measurements, light coherence is needed. This explains the interest to consider monochromatic wave; for such a wave with frequency f, U can be decomposed in spacial and temporal parts as:

$$U(\vec{r}, t) = U(\vec{r}).e^{-i2\pi ft} \tag{2.10}$$

By injecting Eq.2.10 into Eq.2.9, we obtain:

$$\Delta U(\vec{r}) - \frac{4\pi^2 f^2 \eta(\vec{r})^2}{c^2} U(\vec{r}) = 0 \tag{2.11}$$

By introducing the wave vector $k = 2\pi\eta/\lambda$, and the wavelength $\lambda = c/f$, we can state that each component of a monochromatic wave in a linear, isotropic, homogeneous, and nondispersive medium abides by the equation:

$$\Delta U(\vec{r}) + k^2(\vec{r}) U(\vec{r}) = 0 \tag{2.12}$$

Please note that $\eta(\vec{r})$ is the complex refractive index, its real part represents the standard refractive index, used for example in the Fresnel's law [30], while its imaginary parts represents the absorption of the object[119].

Eq.2.12 is known as the Helmholtz equation and it may be difficult to interpret it. Even though we know it contains everything about complex multiple scattering, Snell Descartes refraction law etc, there is no notion of light rays, interfaces, refraction, reflections. Even the simplest concept of propagation direction is not accessible. That is why a further development of the Helmholtz equation is required.

The extended use of monochromatic wave in this chapter will bring us to use its Fourier representation that is more compact. That is why before diving into Helmholtz equation, we will start with reminders of plane waves and Fourier transform.

2.2 PLANE WAVES AND FOURIER TRANSFORM

2.2.1 Plane wave definition

Let consider a monochromatic plane wave $U^{\vec{\mu}}$ of frequency f, with equation:

$$U^{\vec{\mu}}(\vec{r}, t) = e^{2\pi\vec{\mu}.\vec{r}}.e^{-2i\pi ft} = e^{2i\pi(\mu_x x + \mu_y y + \mu_z z - ft)} \tag{2.13}$$

where $\vec{\mu}$ is the propagation direction. Such a wave propagates in a homogeneous medium of refractive index η_0 if its propagation direction verifies:

$$|\vec{\mu}| = \eta_0/\lambda \tag{2.14}$$

as shown in Appendix A.3 and Eq.A.1. From the latter definition, we obtain that λ is the wavelength of a plane wave propagating in the vacuum ($\eta_0 = 1$) and η_0/λ is the wavelength of a plane wave propagating in a medium with optical index η_0.

In the same way that the Helmholtz Eq.2.12 focuses only on the spacial part of the scalar representation of the electromagnetic field, we will now focus on the spacial part of the field of U.

$$U^{\vec{\mu}}(\vec{r}) = e^{2\pi\vec{\mu}.\vec{r}} \tag{2.15}$$

To illustrate the shape of plane waves as defined by Eq.2.15, let first consider the case of a plane wave $U^{\vec{\mu}_0}$ with $\vec{\mu}_0 = (0, 0, \eta_0/\lambda)$ as plotted in Fig.2.1 (amplitude and angle of the complex field at $Z = 0$ its Fourier transform at $Z = 0$ as defined afterwards in section 2.2.2, and its angle a $X = 0$). The figure shows that, for this particular plane wave:

- the amplitude is constant everywhere, there is no attenuation,
- the phase along the Z direction enables us to visualize the wavelength,
- the propagation is along the Z direction,
- if the time was reintroduced by using Eq.2.13, we would find that it propagates along increasing Z.

Let now consider the case of a plane wave $U^{\vec{\mu}_1}$ with none null μ_{1y}, $U^{\vec{\mu}_1} = 1/\sqrt{2}(0, \eta_0/\lambda, \eta_0/\lambda)$. As one may expect, this plane wave is propagating at 45° as shown on Fig.2.2.

A plane wave with wavelength λ propagating in a homogeneous medium, is entirely defined by its propagation direction $\vec{\mu}$. $\vec{\mu}$ is a 3 dimensional vector, but due to the con-

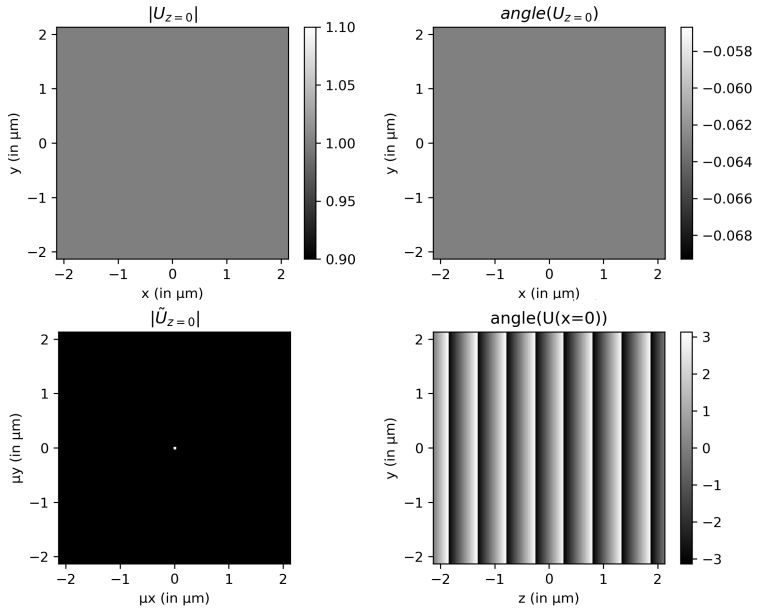


Figure 2.1: Plane wave along Z direction and its Fourier transform

straint of Eq.2.14, it can be characterised by only 2 numbers if we consider that the waves propagates from the negative Z toward the positive Z. As Z axis will be used as the principal direction of light propagation, plane waves will be expressed in function of its X and Y direction: $\vec{\mu} = (\mu_x, \mu_y, \sqrt{\frac{n_o^2}{\lambda^2} - \mu_x^2 - \mu_y^2})$

This formulation helps us to better understand 3D plane wave, since it is parameterized only by two numbers (μ_x, μ_y), it is completely characterized in 2D. In other words, if we know a plane wave on a 2D plane, we have all the information about its 3D behaviour. Afterwards, by expressing any wave as a sum of plane waves (Fourier transform), we will also be able to compute the light field U_z at a plane Z=z from the knowledge of the light field U_0 at position $Z = 0$ where we used the notation:

$$U_z : (x, y) \mapsto U(x, y, z)$$

. These words explain why Fourier transform is at the core of physical optics.

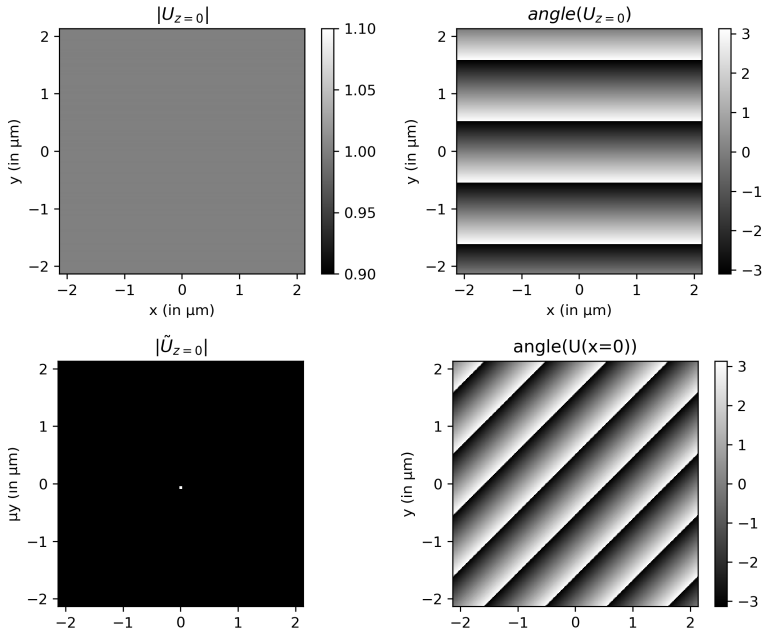


Figure 2.2: Plane wave along Y and Z direction and its Fourier transform

2.2.2 Fourier transform

In this part $(\mu_x, \mu_y) \mapsto \tilde{U_z}(\mu_x, \mu_y) = FT_{2D}(U_z)(\mu_x, \mu_y)$ are two notations for the 2D Fourier transform of $(x, y) \mapsto U_z(x, y)$ defined as:

$$\tilde{U_z}(\mu_x, \mu_y) = \int U_z(x, y) e^{-2i\pi(\mu_x x + \mu_y y)} dx.dy \tag{2.16}$$

With this definition, the inverse formula is:

$$U_z(x, y) = \int \tilde{U_z}(\mu_x, \mu_y) e^{2i\pi(\mu_x x + \mu_y y)} d\mu_x.d\mu_y \tag{2.17}$$

In the same way that we can represent the wave U on the Cartesian axes of a direct orthonormed system X, Y, Z, we can also express it on an other complete, orthonormed system based on Fourier series.

The Fourier transform of U_z can be interpreted as the decomposition of the light field U on a plane waves basis, which is parameterized by coordinates (μ_x, μ_y). For the case of a

plane wave $U^{\vec{\mu}}$, Eq.2.16 shows that the Fourier transform of $U_z^{\vec{\mu}}$ is null for every direction, except its own propagation direction $\vec{\mu}$:

$$FT_{2D}(U_z^{\vec{\mu}_0})(\vec{\mu}') = e^{2i\pi\mu_z z}\delta(\vec{\mu}' - \vec{\mu}) \tag{2.18}$$

where δ is the Dirac function. This situation was illustrated in Fig.2.1 for the case of a plane wave propagating along Z axis and in Fig.2.2 for a plane wave tilted at 45°. In these two cases, we see that the Fourier transforms of U_0 are Diracs, centered on (0,0) in the first case and shifted toward bigger μ_y in the second case.

From here it is clear that every Dirac in the Fourier transform is associated with a planar wave.

2.2.3 Free propagation

In the case of homogeneous medium, it is useful to know how to propagate a wave from one plane to another without having to know the field in every intermediate plane. For example if a detector is placed far away from the simulated object, then it is expensive and useless to simulate the light field in all the volume in between this object and the detector. Other use cases are the multi-slices models, where a simulated object is sliced into a discrete stack of 2d slices. Therefore it is necessary to be able to go from one slice to the next. Such propagation are performed efficiently in the Fourier space.

Let us consider a light field U. We want to compute U_z from a known U_0. Since U satisfies the Helmholtz equation (Eq.2.12), applying the Fourier transform (Eq.2.16) on Eq.2.12, with homogeneous $\eta = \eta_0$ leads to:

$$4\pi^2\left(-\mu_x^2 - \mu_y^2 + \frac{\eta_0^2}{\lambda^2}\right)\tilde{U} + \frac{\partial^2\tilde{U}}{\partial^2 z} = 0 \tag{2.19}$$

It is important to remember that each coefficient of the Fourier space represents a plane wave satisfying the constraint Eq.2.14 which can be re-expressed as $\mu_z^2 = -\mu_x^2 - \mu_y^2 + \eta_0^2/\lambda^2$. By Using that latter equation and by looking for an exponential solution of the

partial derivative equation (PDE) Eq.2.19, *i.e.*: $\tilde{U} = C_1 e^{C_2 z}$ with $C_1(\mu_x, \mu_y)$ and $C_2(\mu_x, \mu_y)$ complex, Eq.2.19 becomes:

$$(2\pi\mu_z)^2 C_1 e^{C_2 z} + C_2^2 C_1 e^{C_2 z} = 0 \tag{2.20}$$

This leads to setting $C_2 = 2i\pi\mu_z$, and since $\tilde{U}(\mu_x, \mu_y, 0) = \tilde{U_0}(\mu_x, \mu_y)$, the value for C_1 is trivial: $C_1 = \tilde{U_0}(\mu_x, \mu_y)$.

Finally the relation between $\tilde{U_0}$ and $\tilde{U_z}$ is a product:

$$\tilde{U_z} = \tilde{U_0}.\tilde{H_z} \tag{2.21}$$

where $\tilde{H_z}$, is known as the spectro angular propagation kernel, defined by:

$$\tilde{H_z} = e^{2i\pi\mu_z z} \text{ with } \mu_z = \sqrt{\frac{n_o^2}{\lambda^2} - \mu_x^2 - \mu_y^2} \tag{2.22}$$

Equivalently, field in the real space can be propagated from one Z-plane to another by using:

$$U_z(x, y) = FT_{2D}^{-1}(FT_{2D}(U_0(x, y)).\tilde{H_z}) \tag{2.23}$$

To obtain this formula no approximation of assumption has been made on top of those used to obtain the Helmholtz equation, the homogeneity of the medium and the hypothesis that the waves propagates from the negative Z toward the positive Z. This means that we have an efficient and correct way to propagate light **as long as the refractive index is homogeneous**. This propagation methods is know as the angular spectrum method. For tilted plane wave, the edges of the simulated volume with scatter light [96], a solution to this problem is known as tilt-transfer, see Appendix A, Section A.2.

In Fourier optics, the Fresnel approximation has been applied earlier to light propagation in free space [7]. The Fresnel kernel is expressed as follow:

$$\tilde{H}_z^{Fresnel} = e^{\frac{2i\pi\eta_0 z}{\lambda}} e^{\frac{-i\pi\lambda z}{\eta_0}(\mu_x^2 + \mu_y^2)} \tag{2.24}$$

This operator is valid only if the size of the wave support (for example the size of an aperture from which we want to find the scattering pattern) is small compared to the propagation distance. An other way to understand what does Fresnel formulation approximate is to see it as an approximation of the spectrum angular kernel for low $(\mu_x^2 + \mu_y^2)$ compared to $\frac{\eta_o^2}{\lambda^2}$. Under this approximation we have that $\sqrt{\frac{\eta_o^2}{\lambda^2} - \mu_x^2 - \mu_y^2} \approx \frac{\eta_o}{\lambda} - \frac{\lambda}{\eta_0}\frac{\mu_x^2+\mu_y^2}{2}$. If we inject it in Eq.2.22 we get Eq.2.24, the Fresnel approximation. If we want a physical interpretation of this approximation ($(\mu_x^2 + \mu_y^2) << \frac{n_o^2}{\lambda^2}$) we have to remember that in the Fourier space each point (μ_x, μ_y) represents a plane wave propagating along $(\mu_x, \mu_y, \sqrt{\frac{n_o^2}{\lambda^2} - \mu_x^2 - \mu_y^2})$. This means that the previous approximation is valid only if the field is mainly composed of waves that are propagating parallel or with a small angle to the Z direction. In most cases with thick biological object and off axis illumination the Fresnel approximation does not hold. That is why we will only use the spectro angular propagator.

2.3 EXTENSIVE EXPLORATION OF THE HELMHOLTZ EQUATION

The Helmholtz equation Eq.2.12 governs the physics of light propagation in free space as we saw in §2.2.3 but also in complex media with non homogeneous optical index $\eta(\vec{r})$. Here, we explain how the varying optical index affects light propagation.

In other words, we want to find the light field U, obtained by illuminating an object defined with optical index $\eta(\vec{r})$ different of η_0 in the space region Ω. We know that U verifies $\Delta U + k^2 U = 0$ and we suppose we also know the illumination, characterized by the field U_{inc} that would exist if the object had the homogeneous optical index η_0.

To find U, we follow a usual strategy which uses the notion of Greens's function and which applies for most partial derivative equations (PDE). The steps are summed up below and developed in the remaining of the current section 2.3.

P0: Original problem: $\Delta U + k^2 U = 0$ can be written as $LU = S = 0$ with $L = \Delta + k^2$. The operator L of the PDE is linear, NOT spatially invariant (because k depends of $\vec{r}$) and has a NULL source terme S.

P1: spatially invariant problem: To solve the original problem P0, we study the twin problem P1, $L_0 G = -\delta$, where L_0 is chosen as spatially invariant and which has a source term, a (minus) Dirac distribution located at $\vec{0}$. The minus of the source term is used for later convenience. Any solution G of the twin problem is called the Green's function. It allows to find the solution f of spatially invariant PDE with a non null source term S, $L_0 f = -S$: $f = S \circledast G$. ($\circledast$ being the convolution operator)

Resolution of the original problem: The original problem P0 is manipulated and reformulated as P1. It allows to give the solution of P0 as a convolution of a source term and the Green's function.

Let develop the concept of Green's function and apply this strategy to Helmholtz equation.

2.3.1 Green's function

As we saw in the previous forewords, the concept of Green's function is used to solve PDEs, such as Eq.2.12 in particular.

2.3.1.1 Green's function in a general case

Let L_0 be a linear and spatially invariant operator. A Green's function G is an impulse response of L_0 to a minus Dirac distribution (located at $\vec{0}$), which is expressed as:

$$L_0 G = -\delta \tag{2.25}$$

The Green's function analytical expression is known for multiple partial differential equations, defined by the formula of L_0. Such expression exists for the wave equation [6], the Schrodinger equation [31], the Poisson equation etc. ...

Since the L_0 is spatially invariant, $G(\vec{r} - \vec{r}_0)$ is the solution of:

$$L_0 G(\vec{r} - \vec{r}_0) = -\delta(\vec{r} - \vec{r}_0) \tag{2.26}$$

Finally since a extended source S may be expressed as $S(\vec{r}) = \int S(\vec{r}_0)\delta(\vec{r} - \vec{r}_0)d\vec{r}_0 = S \circledast \delta$, by linearity of L_0,

$$f = S \circledast G \tag{2.27}$$

is a solution of:

$$L_0 f = -S \tag{2.28}$$

where $\circledast$ stands for the 3D convolution on Ω:

$$(A \circledast B)(\vec{r}) = \int_{\Omega} A(\vec{r'})B(\vec{r} - \vec{r'})dr' \tag{2.29}$$

2.3.1.2 *Spatial and frequency expressions of the Green's function*

In our case of Helmholtz equation in 3D homogeneous medium, the linear operator $\Delta + k_0^2$ is spatially invariant, and its Green's function, G, is the solution of:

$$(\Delta + k_0^2)G = -\delta \tag{2.30}$$

which a solution is known and has the following analytical expression:

$$G(\vec{r}) = \frac{e^{ik_0|\vec{r}|_2}}{4\pi\,|\vec{r}|_2} \tag{2.31}$$

Note that G, defined as a solution of Eq.2.30 is not unique because any solution G_0 of $(\Delta + k_0^2)G_0 = 0$ may be added. $G + G_0$ would still be solution of Eq.2.30. To get unicity for G, additional constraint should be introduced. In our case, it is the physics consideration stating that diffracted light should tend toward 0 at infinity. With such an implicit constraint added to Eq.2.30, Eq.2.31 is unique.

As we will see afterwards (in §2.4), we are interested by the Green's function expressed in the Fourier space in order to perform convolution operations efficiently. The 3D Fourier transform of the Green's function, $\tilde{G}(\mu_x, \mu_y, \mu_z)$, is demonstrated in Appendix A.1. As

formulated in the same way as Vico et al [135], we get:

$$\tilde{G}(\vec{\mu}) = \frac{1}{k_0^2 - 4\pi^2\vec{\mu}^2} \tag{2.32}$$

Another form of the Green's function, expressed with mixed coordinates in real and Fourier space, $\tilde{G}(\mu_x, \mu_y, z)$, will be needed in section §2.5. Since $\tilde{G}(\mu_x, \mu_y, z)$ is the inverse Fourier of Eq.2.32 with respect to μ_z, by using the usual Fourier transform formula: $TF_{1D}[z \mapsto e^{-|a|z}](\mu_z) = \frac{2a}{a^2+4\pi^2\mu_z^2}$, and by taking $a = 2i\pi\sqrt{\eta_0^2/\lambda^2 - \mu_x^2 - \mu_y^2}$, we get:

$$\tilde{G}_z(\mu_x, \mu_y) = \tilde{G}(\mu_x, \mu_y, z) = \frac{i\lambda}{4\pi\eta_0 C} e^{2i\pi z\sqrt{\eta_0^2/\lambda^2 - \mu_x^2 - \mu_y^2}} \tag{2.33}$$

where $C = \frac{\lambda}{\eta_0}\sqrt{\eta_0^2/\lambda^2 - \mu_x^2 - \mu_y^2}$ is the cosine of the angle between Z and the propagation vector the wave $(\mu_x, \mu_y, \mu_z(\mu_x, \mu_y))$ which can propagate inside the medium with the optical index η_0.

We recognize the free propagation kernel as already seen in Eq.2.22, therefore we can write:

$$\tilde{G}_z = \frac{i\lambda}{4\pi\eta_0 C}\tilde{H}_z \tag{2.34}$$

2.3.2 Solution of the Helmholtz equation

Our goal is to express the total field U inside Ω knowing the incident field U_{inc} and the refractive index η. We arbitrary separate the total field in two parts, the incident wave as it would propagate if the refractive index was to be constant U_{inc} and the wave diffracted by the sample U_{dif}.

$$U(\vec{r}) = U_{inc}(\vec{r}) + U_{dif}(\vec{r}) \tag{2.35}$$

2.3.2.1 *Lippmann-Schwinger*

We recall that, for a monochromatic wave, U verifies Eq.2.12:

$$\Delta U + k^2 U = 0 \tag{2.36}$$

whereas the illuminating wave verifies:

$$\Delta U_{inc} + k_0^2 U_{inc} = 0 \tag{2.37}$$

By adding $-k_0^2 U_{inc}$ on both side of Eq.2.36 and subtracting Eq.2.37 we get:

$$\Delta(U - U_{inc}) + k_0^2(U - U_{inc}) = -(k^2 - k_0^2)U \tag{2.38}$$

Using Eq.2.35, Eq.2.38 can be rewritten as:

$$\Delta U_{dif} + k_0^2 U_{dif}(\vec{r}) = -(k^2 - k_0^2)U \tag{2.39}$$

U_{dif} is governed by a homogeneous equation with a source term $(k^2 - k_0^2)U$. $k^2 - k_0^2$ is called the scattering potential. According to §2.3.1.1, U_{dif} can be rewritten with a convolution as:

$$U_{dif} = ((k^2 - k_0^2)U) \circledast G \tag{2.40}$$

Finally, by using Eq.2.35 and Eq.2.40, we get the Lippmann-Schwinger solution of equation Eq.2.12

$$U = U_{inc} + ((k^2 - k_0^2)U) \circledast G \tag{2.41}$$

or fully developed, the expression becomes:

$$U(\vec{r}) = U_{inc}(\vec{r}) + \frac{1}{\lambda^2}\int_{\Omega}(\eta^2(\vec{r}') - \eta_0^2)U(\vec{r}')\frac{e^{ik_0|\vec{r}-\vec{r}'|}}{|\vec{r}-\vec{r}'|}d\vec{r}' \tag{2.42}$$

Since the Green's function is spherical, the geometrical interpretation of Eq.2.41 or Eq.2.42 is that each point of the source term S emits a spherical wave. The sum of all the emitted spherical waves creates the total wave.

One could think that the problem is solved, but since U appears in both the left and right sides of Eq.2.41 or Eq.2.42, U cannot be determined in a simple manner.

2.3.2.2 1st Born approximation

An elementary way to get a first order approximation to compute the Lippmann-Schwinger solution is to assume that the object is weakly scattering. If the scattering potential of the object is small then the diffracted wave will be small compared to the incident wave. This is implemented by replacing U by U_{inc} in Eq.2.40:

$$U_{dif} = \left((k^2 - k_0^2)U_{inc}\right) \circledast G \tag{2.43}$$

This formulation called the "first Born approximation" [43] allows for an efficient way to compute the diffracted field but relies on strong approximations that are often not valid [73]. Indeed, a wave that diffracts in the object will not interact again with the object. The validity of such an approximation has been well studied [69] and it is known that for biological imaging, objects bigger than a few microns produce scattering waves as intense as the incoming wave and cannot be neglected.

2.3.2.3 Rytov approximation

For optically thicker objects, object with a phase-shit introduced by the sample of more than a wavelength, an other approximation is widely used which yields better results [84]. The Rytov approximation [26] is expressed by:

$$U(\vec{r}) = U_{inc}(\vec{r})e^{\psi(\vec{r})} \tag{2.44}$$

with the complex phase $\psi(\vec{r})$ defined by:

$$\psi(\vec{r}) = \frac{U_{dif}(\vec{r})}{U_{inc}(\vec{r})} \tag{2.45}$$

Interestingly, the Born approximation can also be view as a first order development of the exponential inside Eq.2.44. Indeed if $|U_{dif}(\vec{r})| << |U_{inc}(\vec{r})|$ then $|\psi(\vec{r})| << 1$ and since $e^x \approx 1 + x$ if $x << 1$:

$$U(\vec{r}) \approx U_{inc}(\vec{r})(1 + \frac{U_{dif}(\vec{r})}{U_{inc}(\vec{r})}) = U_{inc}(\vec{r}) + U_{dif}(\vec{r}) \tag{2.46}$$

The global approximation for Born and Rytov is the same: $U/U_{inc} \approx 1$. The main difference is that Born is formulated as a sum between the incident wave and the diffracted wave whereas Rytov is formulated as a product with an exponential. Rytov approximation is better than Born approximation for thicker and more scattering object [120] but it still fails to model correctly highly scattering object [120].

2.3.2.4 Iterative resolution of the Lippmann-Schwinger equation

In order to capture all the complexity of multiple scattering sample, the Lippmann-Schwinger Eq.2.40 has to be directly solved. It has no analytical solution in general except for particular geometries such as those describing homogeneous cylindrical objects or homogeneous spherical objects [5].

In the general case, the diffraction can be computed, with a precision as good as desired, by numerical means. U is formulated as the solution of a minimization problem which can be found by a minimization procedure such a the gradient descent method (details in Section 3.1.4).

If we introduce the linear operator L_f associated to the scattering potential $f = k^2 - k_0^2$:

$$L_f : X(\vec{r}) \longrightarrow X(\vec{r}) - \int_\Omega X(\vec{r}).f(\vec{r}).G(\vec{r} - \vec{r'})dr' \tag{2.47}$$

We notice that the Lippmann-Schwinger Eq.2.41 can be expressed as:

$$L_f U - U_{inc} = 0 \tag{2.48}$$

U is the solution of a linear system of high dimension which is usually tackled by using an iterative method. An overview of such the method is the following: we start by taking an initial guess on the total field, U_0. Then, at each iteration n, an error ϵ is computed on how wrong the total field U_n is. From this error, a gradient $\nabla\epsilon$ is computed in order to give information on how variation δU_n of U_n influences ϵ, (see §2.3.2.4.1). From this gradient and the current value of the total field, a step of the optimisation algorithm is performed. The simplest update is a gradient descent where $U_{n+1} = U_n - \alpha\nabla\epsilon$.

2.3.2.4.1 Gradient expression The gradient $\nabla\epsilon$ of a real scalar function $\epsilon(U)$, is a vector which connects variation $\delta\epsilon$ of ϵ to variation δU of its variables. When U is a complex vector, $\nabla\epsilon$ is also a complex vector with the same dimension as U and is defined by:

$$\delta\epsilon = Re(\langle\nabla\epsilon^*, \delta U\rangle) \tag{2.49}$$

In our case, we aim at finding a field U which verifies Eq.2.48 as closely as possible. The easiest way to measure the goodness of U is the squared of Eq.2.48 integrated on all the domain Ω:

$$\epsilon(U) = \frac{1}{2}\int_\Omega (L_f U - U_{inc})^2 = \frac{1}{2}\langle L_f U - U_{inc}, L_f U - U_{inc}\rangle \tag{2.50}$$

With this definition, "The more $\epsilon(U)$ is close to 0, the better U is solution of Eq.2.48.". By using the calculus of Appendix A.4, Eq.A.3 allows to extract the gradient:

$$\nabla\epsilon = L_f^H (L_f U - U_{inc}) \tag{2.51}$$

With:

$$L_f^H : Y(\vec{r}) \longrightarrow Y(\vec{r}) - \int_\Omega Y(\vec{r}).f^*(\vec{r}).G^*(\vec{r}-\vec{r'})dr' \tag{2.52}$$

2.3.2.4.2 Detailed algorithm for forward LS computation Now that we have all the pieces to solve the LS equation, the full reconstruction algorithm is detailed below:

Algorithm 1 LS solve $(\eta, U_{inc}, N_{iters})$

Initialize $U_1 = 0$ and $b_0 = 1$ and $p_0 = p_1 = 0$
for $n_i = 1, 2, \ldots, N_{iters}$ **do**
  $b_{n_i} = \frac{1+\sqrt{1+4b_{n_i-1}^2}}{2}$
  $U_{n_i+1} = U_{n_i} + \frac{b_{n_i-1}-1}{b_{n_i}}(U_{n_i} - U_{n_i-1})$ – Nesterov acceleration
  $\nabla\epsilon = L_f^H(L_f U_{n_i+1} - U_{inc})$ – Compute the gradient toward ϵ
  $U_{n_i+1} = U_{n_i+1} - \sigma\nabla\epsilon$ – Update the field U
end for

For objects that mimic biological samples with variation of refractive index difference bellow 0.1, we found that a good convergence is obtain with only 30 iterations with a gradient descent step $\sigma = 0.3$.

2.4 NUMERICAL SIMULATION WITH THE 3D DISCRETIZED GREEN'S FUNCTION

In the previous chapter we claimed that the Lippmann-Schwinger (LS) equation was solved for low scattering sample, and a more complex iterative process has been presented to solve the LS equation even for highly scattering object. But all of those solution were based on Continuous Convolution Operation (CCO). Even though such convolution (CCO) are convenient for solving ordinary differential equation they are known to be challenging to work with. The main issue is to determine the range of integration and the appropriate integrand for analytical solving. In our case we are lucky since we will not have those issues.

2.4.1 Convolution

Convolution has been used first in 1760s by Euler (L. Euleri, Institutionum Calculi Integralis, vol. 2. Petropoli: Impensis Academiae Imperialis Scientiarum, 1768.), but it was not until the end of the 19th century that it received its first name. At the beginning of the 20th century, it earned a general name, “convolution product" in English, “faltung" in German (G. Doetsch, Theorie und Anwendung der Laplace-Transformation. New York: Dover, 1943. p 157). The lack of interest in convolution (it took almost 2 centuries before it was specifically named) in the 18th and 19th century was balanced by the 20th century with the apparition of the convolution theorem.

The convolution theorems states that, under suitable conditions, the convolution of two functions (or signals) is the inverse Fourier transform of the point wise product of their Fourier transforms. The importance of convolution theorems is that they allow computing the corresponding convolution indirectly through fast Fourier transform (FFT).

2.4.2 Discretization and padding

Instead of working with continuous function and convolution, all the numerical simulation and reconstructions will be performed with discretized field and Fast Fourier Transform.

In order to do the numerical computation, one will have to work on a discrete bounded 3D space. The easiest model is a cubic volume divided uniformly in cubic voxels lying on a regular mesh of size n × n × n. The sides of the voxel are $dv \times dv \times dv$ Let x_{3D}, y_{3D} and z_{3D} be the 3D matrices of the coordinates on a mesh as defined above, centered on (0, 0, 0). Let use the symbols × for the conventional matrix product and . for the pointwise product.

For example, the 3D complex matrix of a unitary plane wave that propagates along the direction $\vec{k} = (k_x, k_y, k_z)$ is:

$$U_{\vec{k}} = e^{i(k_x x_{3D} + k_y y_{3D} + k_z z_{3D})} \tag{2.53}$$

The discretization of the simulated object is straightforward, one just have to take the value of the refractive index at every point of the discretized volume to create the 3D matrix of the refractive index η_{3D}. From which the 3D matrix of the scattering potential is generated by: $f_{3d} = \left(\frac{2\pi}{\lambda}\right)^2(\eta_{3D}^2 - \eta_0^2)$

Because of the singularity of the Green's function (Eq.2.31) and its Fourier transform (i.e. from Eq.2.32, $\tilde{G(\mu)} = \frac{1}{k_0^2 - \mu^2}$), G_{3D} (the 3D matrix of the discretized Green's function) cannot be defined through a naive discretization of $G(\vec{r})$. In the next section, we will describe how G_{3D} has to be defined in order to minimize the approximation error with respect to the continuous model (Eq.2.31).

2.4.2.1 discretization of the Green's function

This section is mainly based on the work of T.-A. Pham et Al [96] and Vico et al. [135].

2.4.2.1.1 Truncated Green's function The refractive index is assumed constant outside of the region of interest Ω. This is convenient in the way that we can work not with the full Green's function but with a truncated version G_t:

$$G_t(\vec{r}) = \text{rect}\left(\frac{\|\vec{r}\|}{2\sqrt{3}L}\right) G(\vec{r}) \tag{2.54}$$

With rect defined as such: $\text{rect}(x) = \{1 \text{ if } \|x\| < 0.5, 0 \text{ otherwsise}\}$, and $L = n * dv$, the size of the volume.

Fig.2.3 is a visual representation in 3D of why the Green's function only have to be considered inside a sphere of radius $\sqrt{3}L$. This truncated Green's function still has a singularity, but not its Fourier transform. That is satisfactory since the convolution will be computed as pointwise product in the Fourier space, That is why we are more interested about getting a correct discretization in the Fourier space than in the spatial domain.

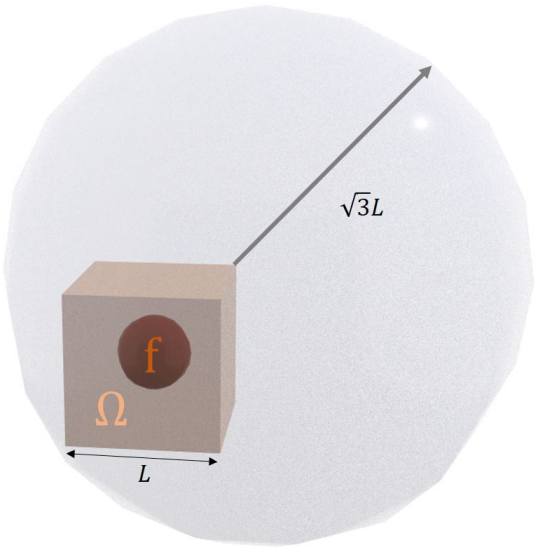


Figure 2.3: Visual representation of the volume used during a convolution with the 3D Green's function and f that has a finite support Ω

Fourier transform of the Green's function $\tilde{G}_t(\vec{r})$ is:

$$\begin{aligned}\tilde{G}_t(\vec{r}) =& \frac{1}{\|\vec{r}\|^2 - k_0^2}\Big(1 - \mathrm{e}^{i\sqrt{3}Lk_0}(\cos(\sqrt{3}L\|\vec{r}\|) \\ &+ ik_0\sqrt{3}L\,\mathrm{sinc}(\sqrt{3}L\|\vec{r}\|)\Big)\Big)\end{aligned} \tag{2.55}$$

The apparent singularity when $\|\vec{r}\|^2 = k_0^2$ is inconsequential. Indeed it can be extended by continuity when $\|\vec{r}\|^2 = k_0^2$ with:

$$\tilde{G}_t(\vec{r}) = i\left(\frac{\sqrt{3}L}{2k_0} - \frac{e^{i\sqrt{3}Lk_0}}{2k_0^2}\sin\left(\sqrt{3}Lk_0\right)\right)$$

Now that $\tilde{G}_t(\vec{r})$ is smooth, its discretization is straightforward, and will be noted $\tilde{G}_{3D}^{4p}$. The ^{4p} is there to point out that a 4 time padding around Ω is required in order to be coherent with our initial support size for $G_t(\vec{r})$ as shown on Fig.2.3. If we want to be more precise, since the support diameter of $G_t(\vec{r})$ is $2\sqrt{3}L \approx 3.4L$, a 3.4 time padding should be sufficient. But it is more commun to work with an integer padding number. A

4 times padding in 3 dimensions means that if we want to simulated a volume Ω with a size $n \times n \times n$, we need to do all the computations (FFT, compute of Green's matrix etc) on matrix of size $4n \times 4n \times 4n$. This matrix is 64 times bigger that its un-padded pair. This means that the memory consumption can become very intense, and limit the simulation capacities. On the computational burden side, what used to takes minutes on an unpadded volumes, now takes hours. Favorably for us, an equivalent truncated Green's function but that only requires a 2 times padding exists.

2.4.2.1.2 Green's function pre-computation for only 2 times padding T.-A. Pham et Al [96] show that using the previous truncated Green's function with a 4 times padding is equivalent as using the following modified Green's function with a 2 times padding:

$$G_{2t}(\vec{r}) = \frac{8}{4^3} \sum_{\vec{s} \in [\![0;1]\!]^3} \mathbf{F}^{-1} \left(\tilde{G}_t([2 \cdot -\vec{s}]) \right) [\vec{r}] \mathrm{e}^{\frac{-i\pi}{\mathrm{n}} \vec{r}.\vec{s}}$$

This modified kernel is smooth just as the truncated Green's function, so the discretization will not be an issue, the discrete modified Green's function notation is: G_{3D}^{2p}. The first thing to notice is that this time the modified Green's function is defined in the spatial domain. To use it inside the convolution theorem, one has to take its Fourier transform, or in the discretized world, its Fast Fourier Transform: $\tilde{G}_{3D}^{2p} = FFT(G_{3D}^{2p})$

Since it is defined on the spatial domain, a graphical interpretation is possible on why adding and translating can reduce the padding requirement (see Fig. 2.4).

The main advantage of using the translated Green's function is to fill more efficiently the simulated space Ω with the simulated object f. Fig.2.4 shows a 2D illustration. Of course, in our case in 3D, it is not 4 but 8 Green's functions that are used to filled a cube instead of 4 circles used to fill a square.

Before detailing the final discretized equation to compute Born or Rytov or LS model, a last way to discretize the Green's function has to be presented. It is not supported by strong mathematical theory but it behaves notably well. The idea for this third Green's

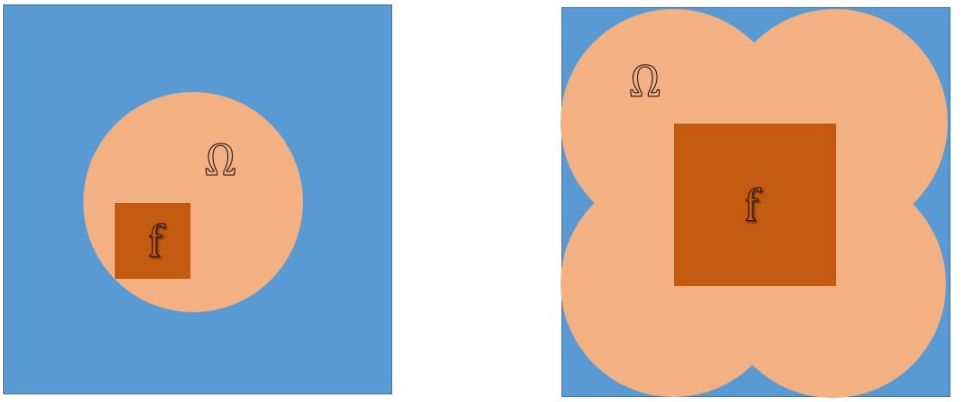


Figure 2.4: Graphical interpretation of the modified Green's function

discretization is to truncates the Green's function numerically.

First a naive discretization of Eq.2.31 is performed, noted $G_{3D}^{n\_ut}$ with "n_ut" for naive un-truncated. We said earlier that we could not discretized it because of its singularity, but most of the time no voxel actually hits the singularity, and even if one voxel goes to high or NaN it can be replaced by any arbitrary "big" value, for example 1e10. Then this matrix goes trough an exponential decay on the edge thus creating our final naive Green's function: G_{3D}^{nt}

After describing 3 way to discretized the Green's function, a naive discretization with an numerical exponential decay on the edge, a truncated Green's function with the $\mathrm{rect}(.)$ function, and a 'P2' truncated Green's function that only needs a 2 times padding.

In order to visualize each Green's function discretization we will used the intensity of it. The intensity of a complex field or a complex kernel is defined as such:

$$I(U) = |U|^2 \tag{2.56}$$

In the Fourier space, the behaviour for $\mu_z = 0$ and $\mu_z = 2\frac{\eta_0}{\lambda}$ frequency of $\tilde{G}_{3D}^{2p}$, $\tilde{G}_{3D}^{4p}$ and $G_{3D}^{n\_ut}$ is shown on Fig.2.5

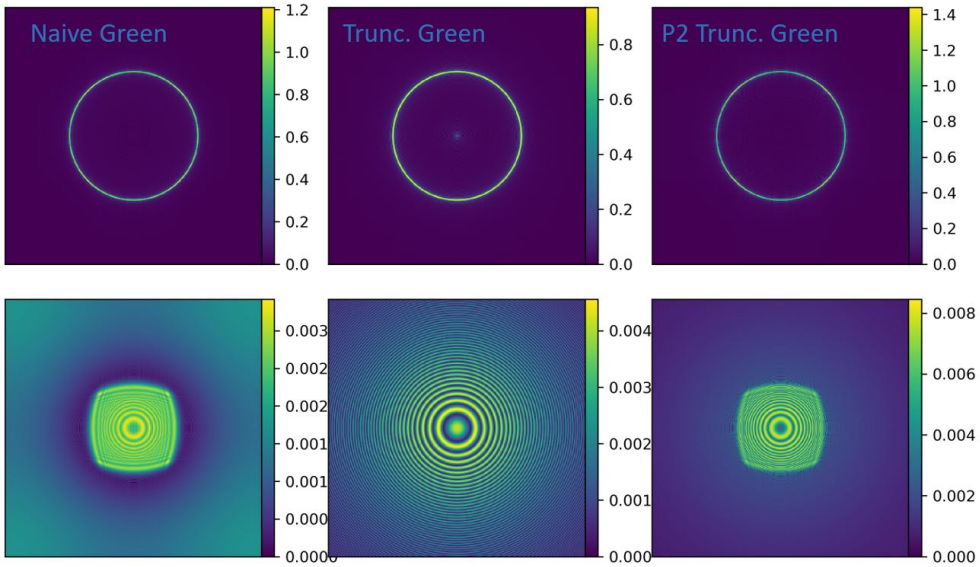


Figure 2.5: Green's function kernel in the Fourier space for 3 different discretizations

2.4.3 Results comparison

Since the First Born and Rytov approximation are based on a low-scattering approximation, we will use a highly scattering sphere to understand the limitations of such model, and the performance of the direct solving of the Lippmann-Schwinger equation. The main advantage of simulating a sphere is that the Mie theory [85] provides an analytical solution against which we can compare the 3 models detailed previously.

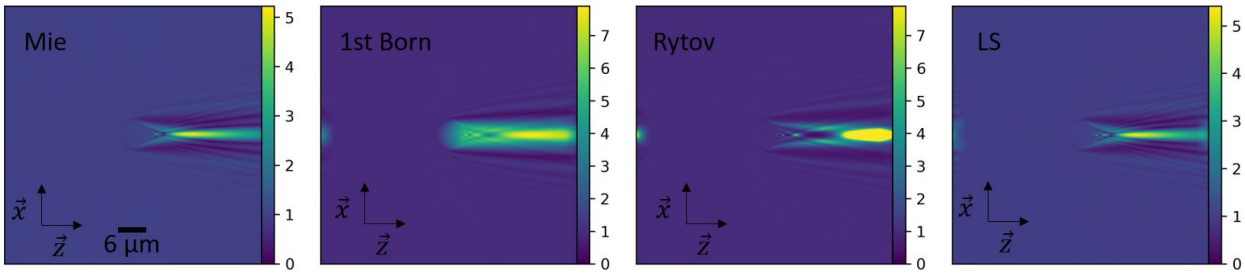


Figure 2.6: Light diffraction by 3 different models of a 3 µm radius sphere of refractive index 1.4 in a 1.33 refractive index medium with the $P2_{trunc}$ Green's function

On Fig.2.6 the First Born approximation fails to model how this highly scattering sphere diffracts light. Even from the left part of the sphere we can see that the field intensity is

too high. This might be related to the fact that the Born approximation does not conserve energy. Indeed the final field has more energy than the incoming field, this extra-energy is found inside and after the sphere.

The Rytov approximation does conserve energy and it looks less wrong, but it does not predicts correctly the destructive interference nor the maximal energy locations.

Finally, and as expected, **the LS models looks close to the MIE theory**, LS is the most accurate model we present in this document. To study the influence of the Green's kernel discretization we will use this model that is well describing the underlying physic.

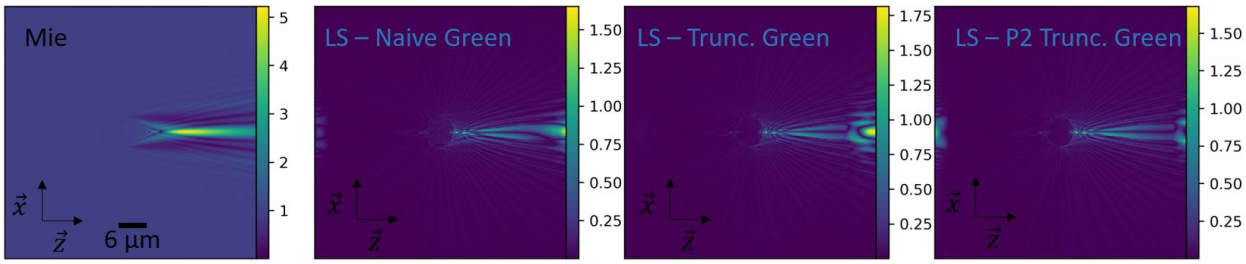


Figure 2.7: Light diffraction by 3 different discretizations of the Green's function of a 3 µm radius sphere of refractive index 1.4 in a 1.33 refractive index medium

On Fig.2.7 the first image is the intensity of the field according to the MIE theory, whereas the other plots are the absolute error between the MIE theory and the LS model with the chosen discretization.

All discretizations behave the same in the center of the field. For the truncated Green's function the error approximation rises higher than the 2 others on the edge of the simulated field. This might explain why a 4 times padding is required with this discretization against only 2 with the P2 pre-computed Green's function. For the naive ED (exponential decay) Green's function no theoretical value for padding exists, but it looks like it behaves comparably to the P2 Truncated Green's function.

As long as the discretization is fine enough (and avoids singularity) the prediction error comes most from the model chosen to model light than the chosen discretization method. The LS model is more accurate than Born or Rytov, but it takes roughly 30 times longer to compute.

2.4.4 Discussion on memory requirement

From the previous section one could think that the LS model in almost perfect and could easily be used to simulate light scattering in the biological objects, like embryos. But in order to image an embryo, we need to simulate a $100 \times 100 \times 100\ \mu m^3$ or bigger volume. Considering the need of using a 2 times padding on top of a $\lambda/4$ discretization in all dimension (according to Fig.2.9 that will be introduced later in section 2.5), we can say that the simulation matrix would have a size of $1500 \times 1500 \times 1500$. This takes up to 27 Gb for one single field stored on double float. If we take into account that at least 6 fields need to be kept in memory (U_{in}, U_{dif}, error, gradient, Green's function, scattering potential), we understand that it is not feasible to model such a big object with this model.

If we consider this time a $10 \times 10 \times 10\ \mu m^3$ volume, everything can be easily done on the GPU since the 6 required fields would only take 0.16Gb if stored on double float.

New model need to be considered in order to simulate wider field of view to image large object.

2.5 MULTI-SLICE MODEL

The main idea behind this section is that merely no reflection occurs in biological samples where optical index varies in a small range (1.33-1.35). Such a knowledge comes from the exact theory of Mie [5] or from Lippmann-Schwinger simulations. Therefore, we will make the approximation that light field completely propagates with increasing Z. So instead of considering the 3D field and its 3D diffraction, we will to compute the field slice after slice.

2.5.1 Multiple Born scattering (MBS)

From Eq.2.23 it is easy to propagate a wave between two planes if the refractive index is constant in between. Now we will detail the calculus to take into account a non-homogeneous refractive index.

According to the Lippman-Schwinger equation Eq.2.41, if $f = k^2 - k_0^2$ is the scattering

potential, after developing the convolution we get:

$$U_Z = U_{incZ} + \int dz \, (f_z U_z) \circledast_{2D} G_{Z-z}$$

where, we use the notation that, for any function with 3D variables g, we call g_z the function of 2D variables such as $g(x, y, z) = g_z(x, y)$. By using the convolution theorem and the Eq.2.34 (Green's function with mixed variables), we obtain:

$$\tilde{U}_Z = \tilde{U}_{incZ} + \frac{i\lambda}{4\pi\eta_0 C} \int dz.TF\left[f_z U_z\right] \tilde{H}_{Z-z}$$

To extend this continuous formulation to a multi-slice one, we need to assume that the field is mainly propagating along the Z direction. Enabling us to replacing the integral that is continuous by a sum over P discrete planes separated by a distance Δz. If we now use the z-discretization convention, where $g_p := g_{p\Delta z}$, the previous equation with $\int dz$ replaced by $\Sigma_p \Delta z$ becomes:

$$\tilde{U}_P = \tilde{U}_{inc_P} + \frac{i\lambda}{4\pi\eta_0 C} \sum_{p=0}^{p=P-1} \Delta z.TF\left[f_p U_p\right] \tilde{H}_{P-p} \tag{2.57}$$

and at plan P+1:

$$\tilde{U}_{P+1} = \tilde{U}_{inc_{P+1}} + \frac{i\lambda}{4\pi\eta_0 C} \sum_{p=0}^{p=P} \Delta z.TF\left[f_p U_p\right] \tilde{H}_{P+1-p} \tag{2.58}$$

U_{inc} is the unperturbed wave propagating inside the medium with homogeneous optical index η_0, so it follows the theory of the free propagation we saw in §2.2.3. We have then: $\tilde{U}_{inc_{P+1}} = \tilde{U}_{inc_P}.\tilde{H}_{\Delta z}$. By looking at Eq.2.23, we see that the same is true for H_Z, we have $\tilde{H}_{P+1} = \tilde{H}_P \tilde{H}_{\Delta z}$. From these remarks, we deduce that:

$$\tilde{U}_{P+1} = (\tilde{U}_P + \frac{i\lambda\Delta z}{4\pi\eta_0 C}.TF\left[f_p.U_P\right]).)\tilde{H}_{\Delta z} \tag{2.59}$$

This multi-slice iterative way to model light scattering was introduced by Michael Chen et Al. [12] in 2020, to the best of my knowledge they arrived to this model by trying to introduce Born diffraction (forward and backward scattering) at each slice in order to improve their previous model. Here we detailed all the approximation required to go from the Helmholtz equation to this multiple Born scattering (MBS) model. The only assumptions that we use are that the refractive index varies slowly in the Z direction and that the field propagates mainly in the Z direction. This is mainly correct if the Z discretization is thin enough, and if the backward scattering is low.

2.5.2 Beam propagation methods (BPM)

The previous model detailed in Eq.2.59, can be simplified again if we consider that the light in mainly propagating with small angle with the Z direction and that the refractive index of the object stays close to the medium refractive index.

Using those 2 assumptions, C can be approximate to be constant, $C \approx 1$ and $f_p = 4\pi^2/\lambda^2(n_p^2 - n_0^2) \approx 8\eta_0\pi^2\delta n_p/\lambda^2$, with $\delta n_p = (n_p - n_0)$, this leads to this recurrent formulation:

$$\tilde{U}_{P+1} \approx \left(\tilde{U}_P + TF\left[\frac{2i\pi\eta_0\delta\eta_P\Delta z}{\lambda}U_P\right]\right)\tilde{H}_{\Delta z}$$

$$\tilde{U}_{P+1} \approx TF\left[U_P\left(1 + \frac{2i\pi\eta_0\delta\eta_P\Delta z}{\lambda}\right)\right]\tilde{H}_{\Delta z}$$

And by using the approximation: $exp(u) \approx 1 + u$, we get the final equation:

$$\tilde{U}_{P+1} \approx TF\left[U_P exp\left(2i\pi\frac{\eta_0\delta\eta_P\Delta z}{\lambda}\right)\right]\tilde{H}_{\Delta z} \tag{2.60}$$

This formulation from Eq.2.60 can be physically interpreted. The field U_{P+1} at the $(P+1)^{th}$ plane is computed from the field U_P at the the P^{th} plane as an optics physicist would expect. The field U_P enters in the medium slice P of thickness Δz and of optical index

η_P. So, the optical length seen by the wave inside this slide is $\eta_P \Delta z$ and the optical path difference (compared with U_{inc}) is $\delta\eta_P \Delta z$. Since the wavelength inside the medium is λ/η_0, we expect a phase shift of $2\pi\eta_0\delta\eta_P\Delta z/\lambda$. Finally this phase-shifted field is propagated Δz further with the spectro angular propagator (with tilt transfer, see A.2). This formulation which takes into account multiple scattering has been around since 1970 [33]. It is normally presented as an independent model, but just as the multiple Born model, everything can be deduced from the Helmholtz equation and some approximation.

Even though the beam propagation methods (BPM, detailed in Eq.2.60) can model multiple scattering they have flaws: BPM is inaccurate when the refractive index does not vary smoothly on the Z axis. Due to the paraxial approximation, predictions lose accuracy for high angle illumination. And finally it does not model backward scattering at all. MBS address the last two problems (backward scattering and high angle of illumination).MBS is also more accurate in low angle illumination. Indeed in the same way that Rytov model is more accurate than first Born, the use of the exponential formulation instead of its first order expansion helps the BPM to deal with optically thicker slices.

2.5.3 Numerical implementation and discretization

Since the two previous models formulations only present kernels that are continuous, the discretization process will be much less complex that with the 3D Green's function.

The direct model computes the output field from a list of N_z slice of arrays of size $N \times N$. We call W_k (as "weight") the variable that stores the variations of refractive index. This corresponds to the discretization of the continuous function $\delta\eta_k$.

The incoming complex field U_0 is a $N \times N$ complex field. We assume that we are far enough from the source for the incoming field to be a plane wave with constant phase and unitary amplitude. Then from the incoming field, all the intermediate fields will be computed recursively. Field U_k at the entrance of the k^{th} slice of the object, is computed as such (. denotes element wise multiplication of matrices).

For the BPM model, where F_k stores the scatterring potential:

$$U_{k+1} = FFT_{2D}^{-1}\left[FFT_{2D}\left[U_k.e^{\frac{2i\pi\eta_0 W_k \Delta z}{\lambda}}\right].e^{2i\pi\sqrt{(\frac{\eta_0}{\lambda})^2-\mu_x^2-\mu_y^2}dz}\right] \tag{2.61}$$

For the MBS model:

$$U_{k+1} = FFT_{2D}^{-1}\left[\left(FFT_{2D}\left[U_k\right] + \frac{i\lambda C\Delta z}{4\pi\eta_0} FFT_{2D}\left[F_k + U_k\right]\right).e^{2i\pi\sqrt{(\frac{\eta_0}{\lambda})^2-\mu_x^2-\mu_y^2}\Delta z}\right] \tag{2.62}$$

U_K is the field at the bottom of the object, a last refocusing step is needed to simulate the imaging system. The final field U_{foc}, focused at a distance z_{foc} from field exiting the domain, follows the next equation:

$$U_{foc} = FFT_{2D}^{-1}\left[FFT_{2D}\left[U_K\right].e^{z_{foc}\left(2i\pi\sqrt{(\frac{n_o}{\lambda})^2-\mu_x^2-\mu_y^2}\right)}\right] \tag{2.63}$$

For a focus in the middle of the object, a negative value of z_{foc} needs to be used.

Now that the models have been introduced, and their full implementation detailed, it is time to quantitatively compare the forward predictions of all the models introduced.

2.5.4 Results comparison

In order to compare quantitatively the accuracy of all the models introduced, a phantom will be studied and simulated.

This phantom is composed of 2 “nucleus" represented by sphere of refractive index 1.37 and radius 1 µm embedded in 4 µm radius bigger sphere of refractive index 1.35. Inside this bigger sphere, two “holes" of refractive index 1.33 and radius 1 µm have been placed on the direct opposite of the first two nuclei. The big sphere is covered with a 1 µm thick layer of refractive index 1.39. A 3D visualisation and sectional plot as detailed on Fig. 2.8

From section 2.4 it is clear that with a proper discretization the LS model is very precise compared to an analytical MIE solution. On the case of our synthetic cell, no analytical solution are available. That is why the LS model will be used as a reference when we

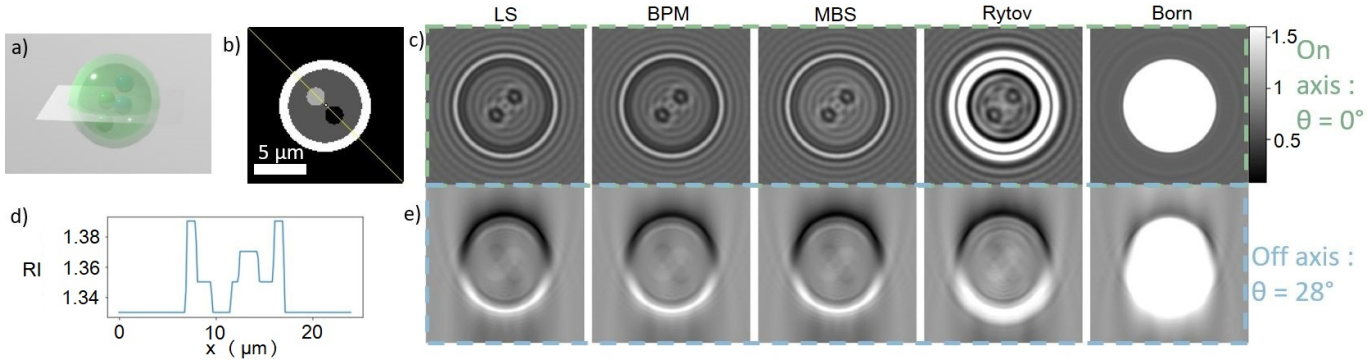


Figure 2.8: Forward simulated image for a highly scattering synthetic cell

simulate the field through our synthetic cell.

First we compare all the model for various discretization size. To demonstrate how the predictions looks like, Fig.2.8 shows the intensity of the focus field for the five models introduced in the previous sections. The source has a wavelength of 532 nm, the medium refractive index is 1.33. The “on axis" lighting is obtained by using a plane wave that propagates only along the Z direction (this means a wave vector of value (0,0,2.5)). The “off axis" lighting is obtain by using a plane wave that propagates with an angle of 28° to the Z direction (this means a wave vector of value (0,-1.1737,2.207))

Fig.2.8 plot has been computed with a discretization size of $dx = \lambda/4$, and a volume of $256 \times 256 \times 256$. Focused at the center slide, the output field has a size of 256×256 with pixel size $dx \times dx$. On top is the simulated image with on axis illumination, on the bottom is the off-axis one. The first thing that appears on this plot is that Born and Rytov are visually “wrong". The low scattering hypothesis does not hold with this highly scattering sample. In order to have a quantitative comparison of all the model, on Fig.2.9 the error of each model for various discretization size is plotted. The error is the mean difference between the intensity simulated with the current discretization and current model, and the intensity of the field obtain with a fine discretization and LS model.

As we observed on the forward plot, both Born and Rytov are applied way out of their validity range. An interesting thing to notice is that for discretization size of $\lambda/2$, all the model based on the discretized 3d Green’s function have the same error. We are

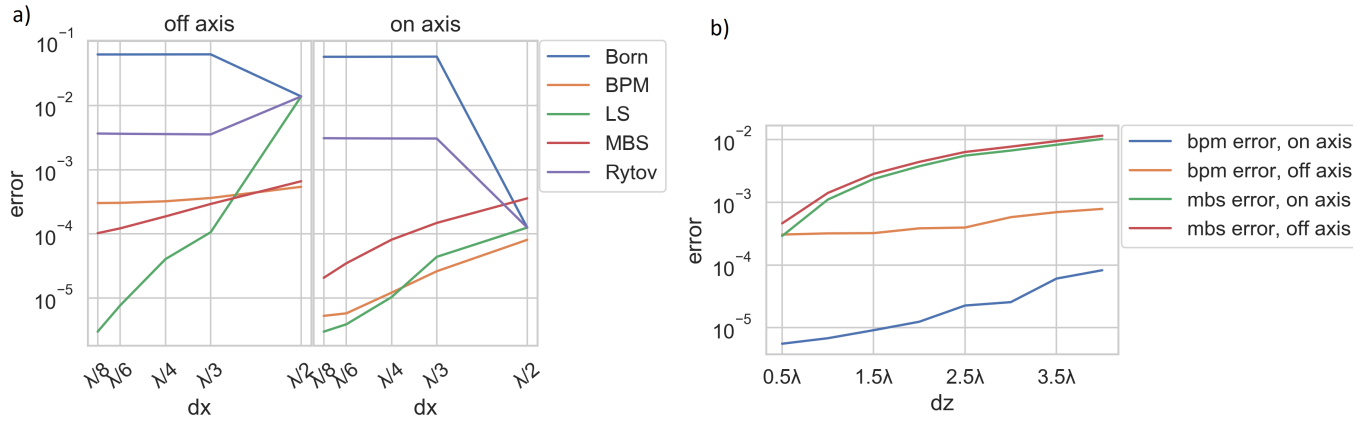


Figure 2.9: Forward model error depending on the discretization size. The object simulated is the same than the one used in Fig.2.8. The error is the mean Squared Error, and the reference is the output of the LS model with thin discretization. a) Error of different forward model respectively to the voxel size. b) Error of multi-slice forward model for various voxel size along the propagation direction: z

outside of the validity range of the discretized Green's function, and then the model used does not matter because all of the error comes from the convolution with a corrupted kernel. Theoretically the MBS model is supposed to be more precise for higher angle, this is indeed true with our highly scattering sample for a discretization size bellow $\lambda/3$. For on axis illumination, the error from the MBS model is almost an order of magnitude higher, regardless of the discretization size.

In order to have a better understanding of the behaviour of the MBS model for high discretization size only on the Z dimension, the Fig.2.9 b) was created. It shows the error for various discretization along the Z dimension, while keeping the discretization size constant in X and Y at $\lambda/4$. The BPM model can maintain a low forward error even if the Z discretization is much higher than the wavelength, up to four wavelength in Figure2.9. This is not the case for the MBS model that becomes worse than BPM as soon as the discretization size become larger than half the wavelength. In the same way that Rytov outperforms Born because it has an "exponential then product" way of computing the diffracted field instead of "linear then sum" for Born. BPM outperforms MBS for high discretization size because if has an "exponential then product" way of computing the diffracted field.

2.6 CONCLUSION

The physics from the Maxwell equation to new multi-slice model of light propagation have been detailed above. Historically, a strong approximation of low-scattering [34] was necessary to predict the diffraction from a 3D refractive object illuminated by oblique incoming light. Nowadays various models have emerged, either based on the discretized 3D Green's function or on multi-slice approach. Each model have its strength and weakness. It isei-ther precise but with high computation cost, or less accurate for high illumination angles but faster and more robust to strong discretization.

Appendix A

Appendix on light scattering physics

A.1 FOURIER TRANSFORM OF THE GREEN'S FUNCTION

In order to express the Fourier transform of the Green's function we can apply the Fourier transform to the definition of the Green's function in Eq.2.31 ($\Delta G + k_0^2 G = -\delta$):

$$FT\left[\Delta G + k_0^2 G)\right] = FT\left[\delta\right]$$

$$(2i\pi\vec{\mu})^2\tilde{G} + k_0^2\tilde{G} = 1$$

Therefore, we get the Fourier transform of the Green's function:

$$\tilde{G}(\vec{\mu}) = \frac{1}{k_0^2 - 4\pi^2\vec{\mu}^2}$$

A.2 TILT TRANSFER

We model off-axis incoming wave has on-axis incoming with a modified spectro angular propagator. This technique is known as "tilt transfer" [96] and avoids diffraction on the border of the simulated fields.

The full detail off this technique is described in detail in [122] (Equation 7 to 11) or in

[96] (Section 3.D).

The "off axis" incoming wave in modelled as an "on axis" incoming wave but the propagation direction is transferred to the propagation kernel [38] by a shift in the Fourier space [39]. The main advantage of this technique is that we provide an incoming field that is continuous even with xy repetition of the field. Therefore when an FFT is performed on the field (an FFT implies that the field is infinitely repeated on each direction), no discontinuity will create scattering artefact. The tilt transfer enables us to diminish the discretization errors and attenuate the scattering artifacts.

A.3 PROPAGATION DIRECTION VECTOR NORM IS CONSTANT FOR PLANE WAVE

From Eq.2.13 we have $U(\vec{r},t) = e^{i2\pi\vec{\mu}.\vec{r}}.e^{-2i\pi ft}$. By injecting this equation inside the Helmholtz Eq.2.12 with homogeneous medium (ie: $\Delta U(\vec{r}) + k^2\eta_0^2 U(\vec{r}) = 0$), we get:

$$-4\pi^2 \left|\vec{\mu}\right|^2 + 4\pi^2 \frac{\eta_0^2}{\lambda^2} = 0$$

And this leads to:

$$|\vec{\mu}| = \frac{\eta_0}{\lambda} \tag{A.1}$$

A.4 GRADIENT OF LIPPMANN-SCHWINGER EQUATION

The definition of the gradient that we use in this document is defined in by the relation:

$$\delta\epsilon = Re(\langle \nabla\epsilon, \delta U \rangle)$$

In our case, the error we will aim to minimize is the squared of Eq.2.48 integrated on all the domain Ω.

$$\epsilon = \frac{1}{2}\int_\Omega (L_f U - U_{inc})^2 = \frac{1}{2}\langle L_f U - U_{inc}, L_f U - U_{inc}\rangle \tag{A.2}$$

From this we can calculate the derivative of ϵ:

$$\delta\epsilon = \frac{1}{2}\langle L_f\delta U, L_f U - U_{inc}\rangle + \frac{1}{2}\langle L_f U - U_{inc}, L_f\delta U\rangle$$

$$\delta\epsilon = Re(\langle L_f U - U_{inc}, L_f\delta U\rangle)$$

$$\delta\epsilon = Re(\langle L_f^H(L_f U - U_{inc}), \delta U\rangle) \tag{A.3}$$

From Eq.A.3, we can directly extract the gradient:

$$\nabla\epsilon = L_f^H(L_f U - U_{inc}) \tag{A.4}$$

The last step is to compute the L_f^H operator, it is defined as such:

$$\langle L_f^H Y, X\rangle = \langle Y, L_f X\rangle \tag{A.5}$$

In our case:

$$\langle Y, L_f X\rangle = \int_\Omega Y^* \left(X - \int_\Omega X f G\right)$$

$$\langle Y, L_f X\rangle = \int_\Omega Y^*(\vec{r})X(\vec{r})\vec{dr} - \int_\Omega\int_\Omega Y^*(\vec{r})X(\vec{r})f(\vec{r})G(\vec{r}-\vec{r'})\vec{dr}\vec{dr'}$$

$$\langle Y, L_f X\rangle = \int_\Omega X(\vec{r})\left(Y^*(\vec{r}) - \int_\Omega Y^*(\vec{r})f(\vec{r})G(\vec{r}-\vec{r'})\vec{dr'}\right)\vec{dr}$$

$$\langle Y, L_f X\rangle = \int_\Omega \left(Y(\vec{r}) - \int_\Omega Y(\vec{r})f^*(\vec{r})G^*(\vec{r}-\vec{r'})\vec{dr'}\right)^* X(\vec{r})\vec{dr}$$

$$\langle Y, L_f X\rangle = \langle Y - \int_\Omega Y f^* G^*, X\rangle$$

From Eq.A.5 we can identify the L_f^H operator:

$$L_f^H : Y(\vec{r}) \longrightarrow Y(\vec{r}) - \int_\Omega Y(\vec{r}).f^*(\vec{r}).G^*(\vec{r} - \vec{r'})dr' \tag{A.6}$$

Chapter 3

3D refractive index reconstruction for time-lapse imaging

The goal of this chapter is to explain how to reconstruct a 3D object from many off axis acquisitions with a LED array microscope. In the previous chapter most of the notation used came from the physical notation. In this chapter that is more optimisation oriented, we will used notation from the applied mathematics world. We start the chapter by giving an overview over the fine-tuning of the regularisation parameters on simulated spherical objects, before moving to living biological samples. This will provide the reader with a sufficient background to correctly understand the more complex scenario of imaging and analysing growing biological specimens.

3.1 OPTIMISATION FOR 3D RECONSTRUCTION

In this chapter a deep understanding of the physics used is not necessary even though it will help to understand when and why an optimisation algorithm may fail to converge.

3.1.1 Optical light diffraction - Summary

The most valuable part of the previous chapter about the light scattering model is that we have a rather simple yet quite accurate model to simulate the images acquired by an intensity only microscope illuminated by a plane wave.

Indeed, the image depends only on 7 parameters:

· λ**: the wavelength of the illumination source**

· μ_{0x}, μ_{0y}**: the illumination wave vector in the Fourier space of the source.**

· η_0**: the refractive index of the surrounding medium**

· $W = \{\underline{N_k}, k \in \{1, ..., n\}\}$**: The 3D matrix of the refractive index of the object** The object to simulate or reconstruct can be represented as a stack of n_z slice $\underline{N_k}$ of complex refractive indexes. $\underline{N_k} = N_k + iK_k$ where N_k is the real part of the refractive index and K_k the imaginary part of the refractive index that corresponds to the absorption of light by the object.

· dz**: the distance between each slice of W**

· Z_{foc}**: The distance between the end of the object and the focal plan**

· NA**: The numerical aperture of the objective**

The simulated image is defined with an iteration through all the slices of the simulated object:

Algorithm 2 $U_{foc} = LightScat(\lambda, \mu_{0x}, \mu_{0y}, \eta_0, W, dz, Z_{foc}, NA)$

Initialize $U_1 = 1$ and $b_0 = 1$ and $p_0 = p_1 = 0$
for $k = 1, 2, \ldots, N_z$ **do**
$U_{k+1} = FFT_{2D}^{-1}(FFT_{2D}(U_k.e^{\frac{2i\pi W_k dz}{\lambda}}).e^{2i\pi\sqrt{(\frac{n_o}{\lambda})^2-(\mu_x-\mu_{0x})^2-(\mu_y-\mu_{0y})^2}dz})$
end for
$U_{foc} = FFT_{2D}^{-1}(p(\mu_x, \mu_y)FFT_{2D}(U_{N_z+1}).e^{z_{foc}CONJ(2i\pi\sqrt{(\frac{n_o}{\lambda})^2-(\mu_x-\mu_{0x})^2-(\mu_y-\mu_{0y})^2})})$

Where the pupil is defined by:

$$p(\mu_x, \mu_y) = \begin{cases} 0, & \text{if} \sqrt{\mu_x^2 + \mu_y^2} > \frac{NA}{\lambda} \\ 1, & \text{otherwise} \end{cases} \tag{3.1}$$

3.1.2 Reconstruction of an object as an inverse problem

Considering that we can model our physical measurement system as a function f, we want to find W that produced the measured y where $y = f(W)$

3.1.3 Reconstruction of an object as a minimisation problem

As explained in the introduction, we will focus on methods in which only one intensity image is recorded at each illumination angle. Even-though half the information is lost when an intensity image is recorded (the light phase is lost), it is still possible to retrieve the refractive index information of the imaged object. It was first shown when a microscope lamp was replaced by a LED array [142] [141]. Then a lot of work was done to increase resolution and speed in 2D [129] [132] [130]. The 3D volumes were also reconstructed with the same kind of computational microscope [133] [131] [51].

We can consider that when the recording is done we have N_{img} intensity diffraction images I_i under N_{img} different illuminations angles. From this set of measurements, our goal is to find the object W_{recons} that produced the acquired intensity diffraction images.

If our reconstruction algorithm was perfect, the output of our light-scattering model $LightScat(\lambda, \mu_{0x}, \mu_{0y}, \eta_0, W_{recons}, dz, Z_{foc}, NA)$ from the reconstructed object W_{recons} would produce exactly the same image as the recorded image.

$$W_{recons} = \arg\min_{W} \sum_{i=1}^{N_{img}} \left| |LightScat(\mu_{ix}, \mu_{iy}, W)| - \sqrt{I_i} \right|^2 \tag{3.2}$$

Please note that the intensity acquired by a camera is proportional to the square of the field computed with our model $LightScat$.

When we want to address the problem of reconstruction of the 3D map of a sample

for multiple images taken with various illumination angles, a lot of parameters from Algorithm 1 can be considered known and constant. For a given illumination source at a given place, the wavelength λ is given by the manufacturer and the illumination angle (μ_{0x}, μ_{0y}) can be computed with some basic Trigonometry. Even if we don't know the refractive index of the object we are trying to image (W), we know the refractive index in which it is embedded, η_0. The numerical aperture is given by the objective chosen, and dz and Z_{foc} depend of the discretization chosen for W

3.1.4 First gradient descent

3.1.4.1 Without regularisation

To solve our inverse problem formulated as a minimisation problem, a first naive idea is to do a gradient descent algorithm on the data fidelity term ϵ:

$$\epsilon(W) = \sum_{i=1}^{N_{img}} \left| |LightScat(\mu_{ix}, \mu_{iy}, W)| - \sqrt{I_i} \right|^2 \tag{3.3}$$

A gradient descent algorithm is a first order optimization algorithm used to find the local minimum of a function that can be differentiated. It works by doing a steep toward the opposite of the gradient (*i.e.* the steepest descent) and then compute the gradient at this new location and repeat iteratively this process. This idea is formalized in the following algorithm with ϵ the differentiable function, η the gradient step and N_{it} the number of iterations.

Algorithm 3 $W_{opt} = GD(\epsilon, \eta, N_{it})$

Initialize $W_1 = 0$
for $k = 1, 2, \ldots, N_{it}$ **do**
  $\epsilon_k = \epsilon(W_k)$
  $\Delta_{W\,k} = \frac{\partial \epsilon(W_k)}{\partial W_k}$
  $W_{k+1} = W_k - \eta \Delta_{W\,k}$
end for
$W_{opt} = W_{N_{it}+1}$

To have a better idea of the convergence rate and the quality of the reconstructed objects with this algorithm, we will start by considering a simple bead and simulate the acquisition in the same conditions than our time-lapse prototype presented in section 3.2.1.

The object is a 2 µm spherical bead of refractive index 1.37 in a medium of refractive index 1.33. It creates a refractive index mismatch of 0.04. The acquisitions are composed of 84 images of size 256×256 with a pixel size of $0.126 \times 0.126\ \mu m^2$ and a NA of 0.4. The simulated angles range from 0° to 15°. The convergence of the gradient descent algorithm 3 for the first 100 iterations are plotted in Fig.3.1

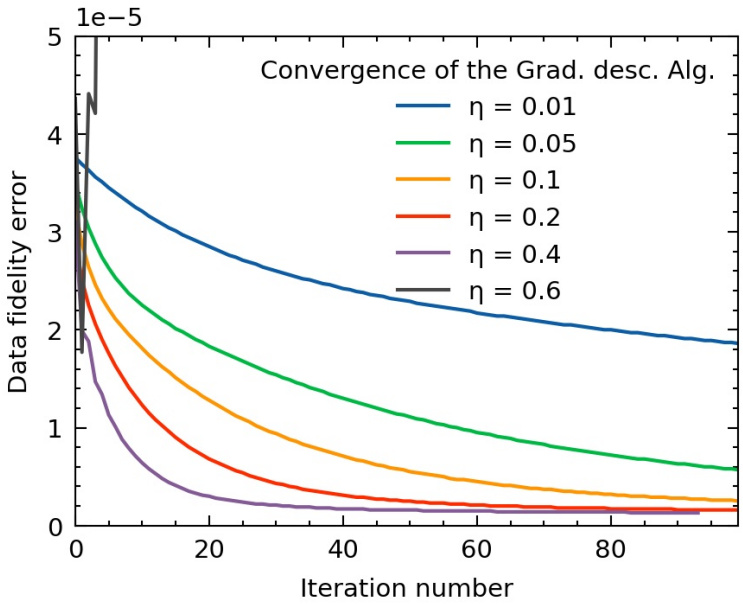


Figure 3.1: Convergence of the data fidelity error for various η using the GD algorithm

For high learning rate (η = 0.2 or η = 0.4) the GD algorithm converges toward the same object. For lower learning rate (η < 0.2) 100 iterations are not enough to achieve convergence. The higher the learning rate the faster the data fidelity decreases. The issue when using learning rates that are too high is the divergence issue. Indeed, when using η = 0.6 the data fidelity error decreases only for the first iterations, then the error starts rising again, like the black line in Fig. 3.1.

However, the equality of the error for different η does not mean that the reconstructed

objects are identical. Fig.3.2 depicts the xy and ZY cross sections for different values of the learning rate η, showing that the GD algorithm converges toward the same object for sufficiently high learning rate ($\eta > 0.1$). This object is a local minimum, indeed once the gradient descent has reached this object or does not move because we can see the error stagnating. This means that the gradient is null or almost null, but the reconstructed object is not the correct spherical object.

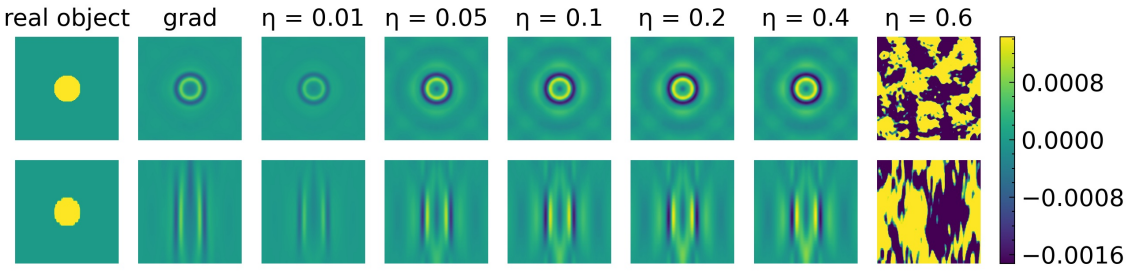


Figure 3.2: xy and ZY slice of the reconstructed 3D bead from the simulated measurement

We can observe that regardless of the hyperparameter η used, only the edges of the object are reconstructed. The center part looks constant and null. This was expected because in our scattering model we consider only the intensity of the light recorded and not its phase. But the intensity of the light only contains information about the gradient of the refractive index (changes). If the refractive index is constant it will not scatter, thus we will have no information about it on our intensity only image. That is why the high frequencies of the object are easier to be reconstructed, contrary to the low frequencies where the information available is much lower.

A second reconstruction artefact is the elongation of shapes along the optical axis (z axis). This is slightly improved as the algorithm converges but the object is still 3 times longer in the z direction than the correct object in this example. This is mostly due to the limited angle coverage and our low NA used.

The first reconstruction artefact (i.e. lack of low frequency of objects) can be dealt with by adding some priors on the object we reconstruct. A first idea would be to impose a non-negativity constraint as we know that biological objects ($W = n - n_0$) are only composed

of positive values. Indeed their refractive index (n) is always higher than the surrounding medium (n_0). A second idea would be to use the fact that most of the biological object we aim to reconstruct are continuous and full object. In this case a Total-Variation (TV) priors that penalises the gradient of the object would help a lot to reconstruct the low frequency since constant pattern would not be penalized by the TV priors but would increase the data-fidelity since the object shape is more correct.

3.1.4.2 With regularisation, positivity

To help the optimisation algorithm to converge towards a correct object, we have to use some prior knowledge about the object we want to reconstruct. We have to keep in mind that the end goal is to image biological samples and that our priors must be biologically compatible.

Since we have a different prior on the common refractive index (the imaginary part of W), and the absorption (the real part of W) we will use the following notation to clarify the next sections:

- W_n will be the usual refractive index, $W_n = real(W)$
- W_{abs} will be the absorption part of the complex refractive index, $W_{abs} = imag(W)$

The positivity is the first prior we will investigate in this section. Positivity means that our object has a higher refractive index than the surrounding water or medium. Indeed the dry mass of a biological sample is correlated to the increase of its refractive index integrated over the whole sample.

To solve our inverse problem formulated as a minimisation problem with a regularisation term $reg_{nneg}(W_n) = nneg * ||ReLU(-W_n)||_2^2$ we need to minimize $\epsilon_{nneg}(W)$:

$$\epsilon_{nneg}(W) = \epsilon(W) + reg_{nneg}(W_n) \tag{3.4}$$

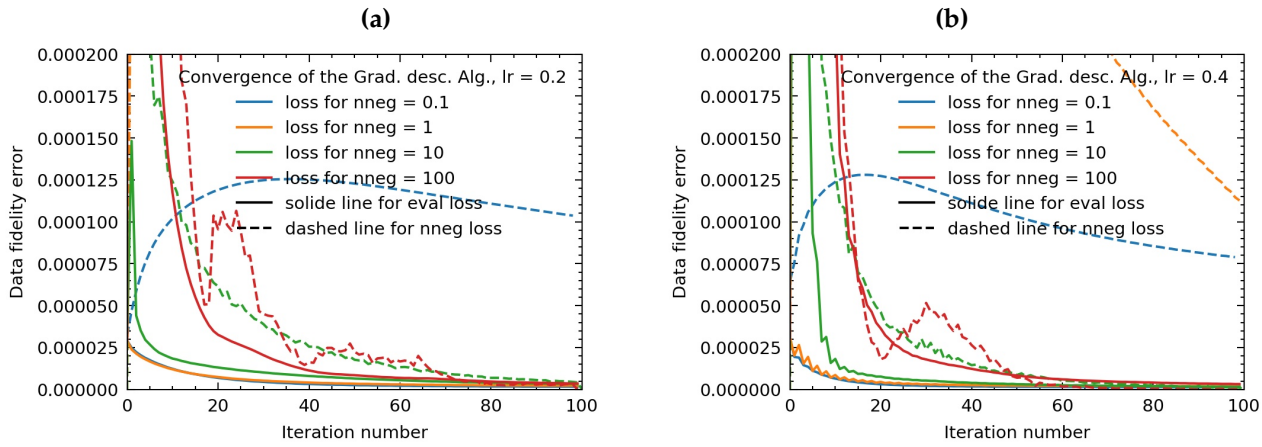


Figure 3.3: Convergence of the data fidelity error and the regularisation loss for $\eta = 0.2$ and $\eta = 0.4$ and various nneg values using the GD_{nneg} algorithm

With $ReLU(x) = \begin{cases} 0 & \text{for} \quad x < 0 \\ x & \text{for} \quad x \geq 0 \end{cases}$

An explicit formulation of the quantity to minimize is:

$$\epsilon_{nneg}(W) = \sum_{i=1}^{N_{img}} \left| |LightScat(\mu_{ix}, \mu_{iy}, W)| - \sqrt{I_i} \right|^2 + nneg * ReLU(-W_n) \tag{3.5}$$

The same naive gradient descent algorithm will be used to evaluate the impact of this regularisation:

Algorithm 4 $W_{opt} = GD_{nneg}(\epsilon, \eta, N_{it})$

Initialize $W_1 = 0$
for $k = 1, 2, \ldots, N_{it}$ **do**
  $\epsilon_k = \epsilon_{nneg}(W_k)$
  $\Delta_{W\,k} = \frac{\partial \epsilon(W_k)}{\partial W_k}$
  $W_{k+1} = W_k - \eta \Delta_{W\,k}$
end for
$W_{opt} = W_{N_{it}+1}$

For very small value of the $nneg$ parameter this problem is almost identical to the minimisation problem without regularisation, and the convergence plot is identical. To avoid

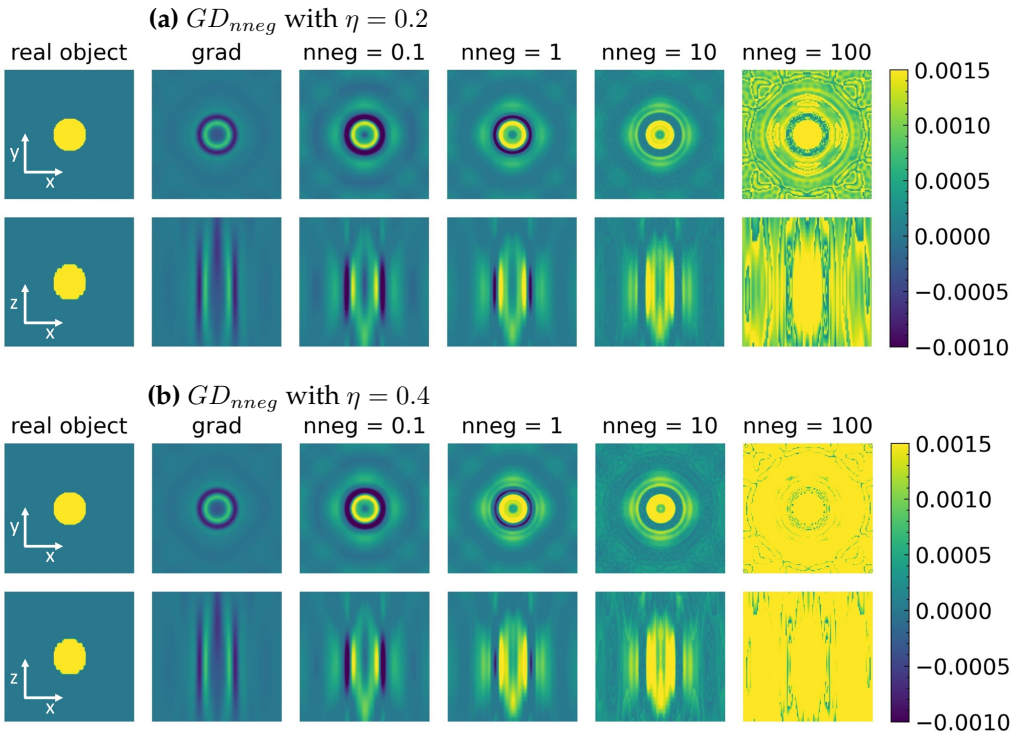


Figure 3.4: xy and xz slice of the reconstructed sphere for $\eta = 0.2$ and $\eta = 0.4$ and various nneg values using the GD_{nneg} algorithm

having to deal with too many hyper-parameters, we will only consider power of 10 for the nneg regularization and a gradient step of 0.2 and 0.4 that provided the faster convergence in the first naive gradient descent. The data fidelity loss and the regularisation loss are plotted in Fig.3.3.

For higher values of $nneg$, the data fidelity error decreases slower than previously, see the red and green solid lines on Fig.3.3. We can see that for $\eta = 0.2$ the loss decreases in a smother way than for $\eta = 0.4$, where it starts oscillating for the green and orange solid lines. If the oscillations are too high the algorithm is not stable and it diverges; this happens for $\eta = 0.6$.

As shown previously, the value of η does not have a large impact on the reconstructed object (comparing Fig.3.4(a) and Fig.3.4(b)) as long as it is high enough for the gradient descent to converge and small enough to prevent divergence of the algorithm.

On the contrary, the $nneg$ parameter has a huge impact on the reconstructed object. First we see that the higher $nneg$, the less negative voxel are present in the reconstruction. This was the main goal of the introduced regularization. When the regularization is too strong, the algorithm converges toward a minimum where all the background has a high refractive index value. This has a good data agreement, but it is far from what we expected. We can also see that the elongation issue along the z direction has not been corrected at all, regardless of the $nneg$ value.

If the non-negativity constraint cannot help with the elongation issue, other regularisations must be considered.

3.1.4.3 With regularisation, positivity and total-variation

The total-variation (TV) norm on a 3D continuous object f is defined as such:

$$TV(f) = \int |\nabla f(\vec{r})| \tag{3.6}$$

The idea behind this regularisation is to penalise noisy objects, as only the gradient of the object is penalized and not its value. In other words, every times that the value of the object increases or decreases the regularisation value increases, but when the value is constant the regularisation is null, no matter how high the value is.

Total-variation minimisation has been used in image processing for image denoising with edge preservation [106] [136]. It has also been suggested as a regularization function in Bayesian reconstructions [92] [95] [94] and used in tomographic scans [118] before being adapted to the optical diffraction tomography field [68].

The formulation on a discrete object W is an isotropic formulation that does not favor vertical or horizontal patterns:

$$TV(W) = \sum_{i=1}^{N_x}\sum_{i=1}^{N_y}\sum_{i=1}^{N_z}|(W[i+1,j,k] - W[i,j,k])^2 + (W[i,j+1,k] - W[i,j,k])^2 + (W[i,j,k+1] - W[i,j,k])^2| \tag{3.7}$$

Other discrete total-variation formulations have been proposed in 2D for images [19] but have not been tested on 3D volumes, yet.

To solve our inverse problem formulated as a minimisation problem with a regularisation term $reg_{nneg,\beta}(W_n) = nneg * ReLU(-W_n) + \beta * TV(W_n)$ we need to minimize $\epsilon_{nneg,\beta}(W)$:

$$\epsilon_{nneg,\beta}(W) = \epsilon(W) + reg_{nneg,\beta}(W_n) \tag{3.8}$$

An explicit formulation of the quantity to minimize is:

$$\epsilon_{nneg,\beta}(W) = \sum_{i=1}^{N_{img}} \left||LightScat(\mu_{ix}, \mu_{iy}, W)| - \sqrt{I_i}\right|^2 + nneg \times ReLU(-W_n) + \beta \times TV(W_n) \tag{3.9}$$

The same naive gradient descent algorithm is used to evaluate the impact of this double regularisation.

The TV regularization was mainly introduced to control the noise level in various inverse-problems on image [107] [125]. On our simple example based on simulation measures, there is almost no noise and that is why the main benefit we were expecting from the TV regularization is its ability to penalise less objects that are full than edges-only object (empty shell).

Looking at the xy slice on Fig.3.5 (a), we can see that the central slice is not filled and a hole in the refractive index value is still present at the center of the sphere. This hole is smaller than without TV regularization, but still present. What was not expected is that

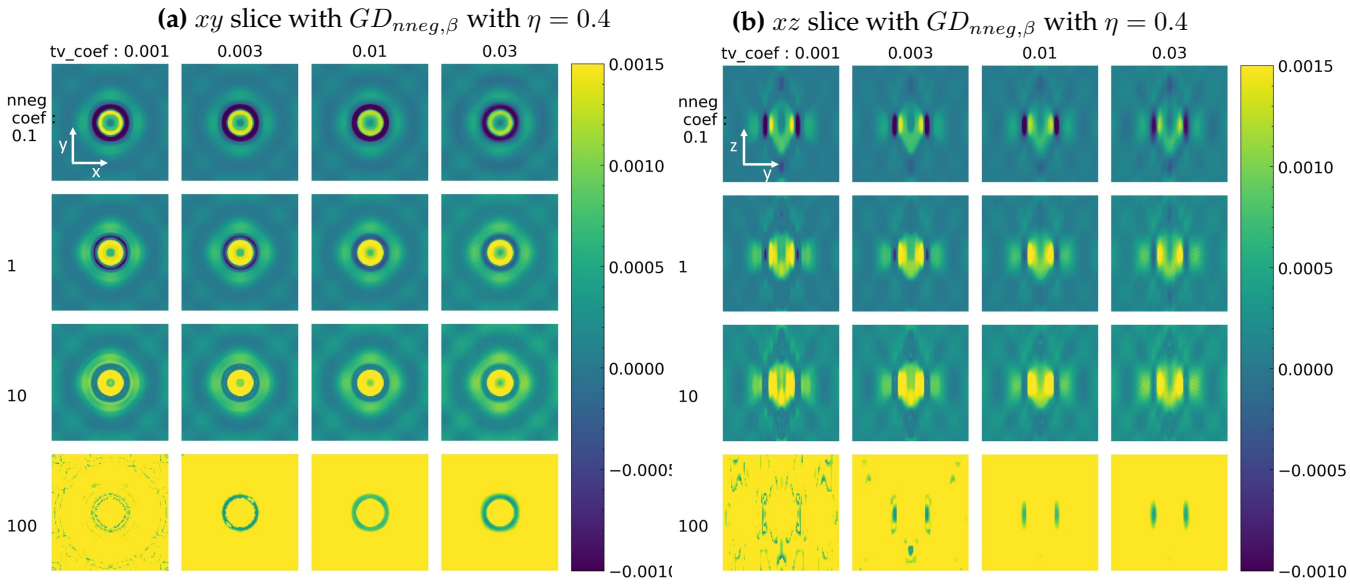


Figure 3.5: xy and xz slice of the reconstructed sphere for $\eta = 0.4$ with various nneg and β values using the GD_{nneg} algorithm

this TV regularisation also decreased the elongation along the Z direction, as it can be observed on the xz slice in Fig.3.5.

This preliminary study with un-optimal optimisation algorithm was performed to give us an insight over which regularization can help with the various problems we are facing in intensity only holography tomography.

3.1.5 Accelerate gradient descent

State of the art for non-convex optimisation on light multiple scattering model is stochastic proximal-gradient algorithm. It was successfully used for high resolution optical tomographic imaging system [59] as well as intensity only imaging system [16].

For standard non-convex optimisation many other approach exists, in the 2D physical optics reconstruction, ADMM [86] and Conjugate gradient methods [45] have been used.

We will first review a conjugate gradient method before comparing it to state of the art proximal-gradient method.

3.1.5.1 *conjugate gradient method*

Conjugate gradient has been first introduced to solve symmetric positive linear system.

For example if A is a real, symmetric, and positive-definite model (as in our case), B is our data and X our parameter. Conjugate gradient is an algorithm to find the unique X_* that solves $AX = B$.

The iterative conjugate gradient can be seen as a regular gradient descent with a conjugation constraint on the direction to follow. An efficient way to enforce it is by forcing the next direction to be built from the current gradient as well as all the previous directions followed.

A non-linear version of the conjugate gradient exists and is suitable to our problem. It has been used in ODT [32] [65], magnetic resonance imaging (MRI) [66], as well as X-ray imaging [101] or spherical tomography [48].

The following algorithms details the conjugate gradient steps for N_{it} iterations to minimize the function ϵ of W:

Algorithm 5 $W_{opt} = CGD(\epsilon, N_{it})$

Initialize $W_1 = 0$
$r_1 = \frac{\partial \epsilon(W_1)}{\partial W_1}$
$d_1 = r_1$
for $k = 1, 2, \ldots, N_{it}$ **do**
  $\alpha_k = \arg\min_\alpha \epsilon(W_k + \alpha d_k)$
  $W_{k+1} = W_k + \alpha_k d_k$
  $r_{k+1} = -\frac{\partial \epsilon(W_{k+1})}{\partial W_{k+1}}$
  $b_{k+1} = \frac{r_{k+1}^T (r_{k+1} - r_k)}{-d_k^T (r_{k+1} - r_k)}$
  $d_{k+1} = r_{k+1} + b_{k+1} d_k$
end for
$W_{opt} = W_{N_{it}+1}$

There are three main differences between the linear and non-linear conjugate gradient descent. First, the residual is always the opposite of the gradient (*i.e.* $r_{k+1} = -\frac{\partial \epsilon(W_{k+1})}{\partial W_{k+1}}$). Second, the distance α_k to move in the new direction d_k is more complex to compute, but any value that truly estimates $\arg\min_\alpha \epsilon(W_k + \alpha d_k)$ can be used. Finally, the formulation

for b_{k+1} mentioned here was derived from the Hestenes-Stiefel paper [47]. Other solutions have been proposed, such as the Fletcher–Reeves [35] $b_{k+1} = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$, the Polak–Ribière [100] or the Dai-Yuan [22] equations. All those variants are equivalent in the linear case and thus produce the same value for b_{k+1}. For non-linear cases, like our inverse problem with the non-linear forward model $LightScat$, there is no guarantee that one expression may provide faster convergence than another.

This non-linear conjugate gradient descent comes with more unknown on the convergence speed than its linear version: $i)$ the more f is different to a quadratic function, the faster the directions d_k will lose their conjugation; $ii)$ if f has local minima we have no assurance that the CGD will converge to the global minimum instead of the closest local minimum; $iii)$ the CGD can only generate n conjugate vectors in a n dimensional space, thus the algorithms is logically restarted every n iterations. This issue does not concern our problematics, as we work in a dimensional space with millions of dimensions, therefore expecting the algorithm to converge in less than a few hundred iterations.

In order to assess the performances of this algorithm we will use the same optimisation problem as in the previous section.

The inverse problem will be regularized by 4 parameters $nneg$,β,$spars$,$spars_{abs}$. "$spars$" stands for "sparsity".

$$\epsilon_{nneg,\beta,l1,l1_{abs}}(W) = \sum_{i=1}^{N_{img}} \left| |LightScat(\mu_{ix}, \mu_{iy}, W)| - \sqrt{I_i} \right|^2 + nneg.ReLU(-W_n) + \beta.TV(W) + l1.|W_n| + l1_{abs}.|W_{abs}|$$

To evaluate this more complex optimisation algorithm the same object will be used as in section 3.1.3: a 2 µm bead of refractive index 1.34 embedded in water (RI = 1.33). 84 intensity images were simulated at angles corresponding to our prototype real data acquisition. We used the conjugate gradient algorithm to solved the regularised problem

from the simulated intensity images.

A grid search on the parameters gave us some base values that produce a good 3D reconstruction in terms of RI shape, noise and value. The parameters that gave the best overall result are: $\beta_{man} = 0.04$, $nneg_{man} = 0.1$, $l1_{man} = 0.01$, $l1_{abs_{man}} = 0.01$. **To have a better understanding of the effect of each parameter we will tune it from zero up to 32 times the initial value. The reconstructed object will change from not regularized enough to overly regularized.**

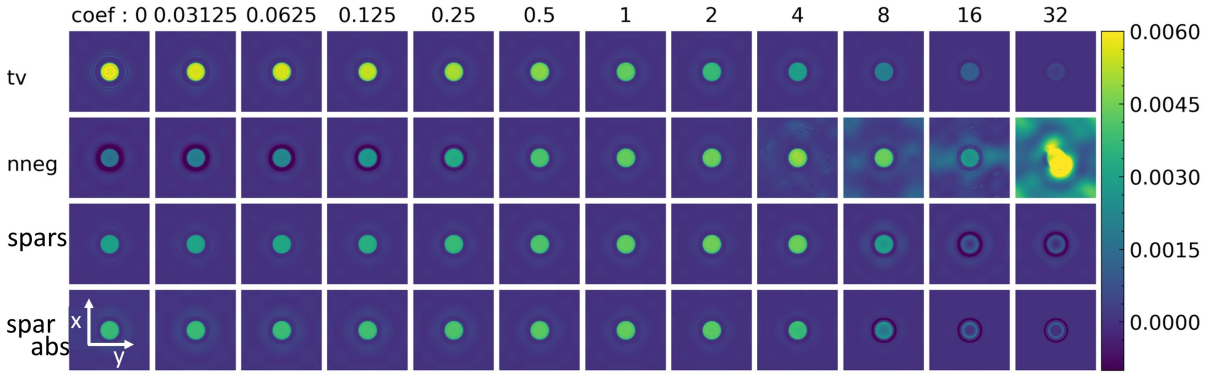


Figure 3.6: xy slice of the reconstructed 3D bead from the simulated measurement. The coefficients in the first column were progressively increased from zero up to 32 times the initial value.

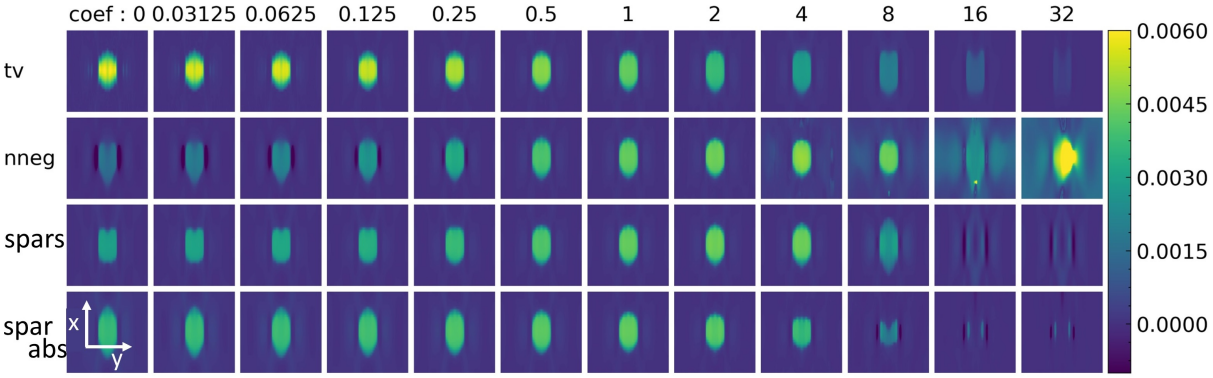


Figure 3.7: yz slice of the reconstructed 3D bead from the simulated measurement. The coefficients in the first column were progressively increased from zero up to 32 times the initial value.

Each regularisation parameter introduced in Section 3.1.4.3 is important to produce

correct reconstructions, as shown in Fig.3.6 and 3.7.

We will review here-below the effect of each parameter:

- tv: the total-variation regularisation tends to smooth out the object in constant patches. If it is set too high, the object gets completely removed, similarly to the sparsity. If the opposite happens, then the object tends to be very noisy and some artefacts appears. Even in this simulation where no noise was added (only some blurring), rings around the reconstructed object and other high-spatial frequency features appear at low tv coefficients. These features are not present in the real object and therefore considered as errors of the regulation method.

- $nneg$: if the non-negativity constraint is too low, a ring of negative RI gap appears around the sphere. This ring acts by decreasing the RI value inside the sphere as well, behaving as low-frequency features (the sphere has constant RI). If we had warranty on the convergence of our optimisation algorithm, with any "high" value of the non negativity constraint, no more voxels with negative RI value should remain, putting the non-negativity loss to zero. This is not the case here: when the $nneg$ parameters increases, the algorithm starts diverging and produces reconstructions that clearly diverge from the real spherical object (see coef 8,16,32 in the nneg line in Fig.3.6,3.7).

- $spars$: the sparsity on the real refractive index tends to flatten the background back to zero. If this paramareter is set too low, the refractive index that should normally be inside the sphere is spread across the whole volume. If it set too high, the object is also considered as background and the whole volume is reduced to zero.

- $spars_{abs}$: the sparsity on the absorption object (imaginary part of the refractive index) is here to ensure that most of the information in stored into the real refractive index part, since we know that the observed objects will be mainly transparent. Interestingly, the larger the regularisation on the absorption, the most compact the reconstructed sphere is (there is less elongation with a higher $spars_{abs}$). Similarly to

the $nneg$, the sparsity in the absorption could theoretically be as big as we want and it should always converge toward a real-only refractive index object. Here again, the algorithm cannot converge if the $spars_{abs}$ is set too high.

Regardless of the parameters used, the sphere is always elongated along the optical axis and this is problematic for two reasons: first, the geometry is not correct and it lowers the z resolution of the microscope; second, the RI distribution is also elongated over the optical axis, thus decreasing the real refractive index values of the reconstructed object. Indeed, the most important part of the scattering fields comes from the optical path difference of the object, and for a given optical path difference if an object is 2 time longer, its refractive index will be 2 time smaller. That is why on Fig.3.6,3.7 the maximum RI gap is 0.006 instead of the simulated 0.010.

Since elongation comes from the intrinsic physical lacks of information, our system will never be quantitative for each voxel value. But it can be quantitative in the sense that the optical path difference (OPD) can be correct, as shown in Fig.3.8. We can compute the optical path difference (OPD) by integrating the distance times the refractive index gap. The OPD value represents the phase shift introduced by the object; if this value is constant over the object and equal to the wavelength (or a multiple of the wavelength), then on a phase sensor one should record a constant phase.

$$OPD((W_k)_{k\in 1...N_z}) = \sum_{k=1}^{N_z} W_k \delta z \tag{3.10}$$

For a $2\ \mu m$ radius sphere with a refractive index gap of 0.01, the maximum OPD is $0.04\ \mu m$, a value that matches the various OPDs from Fig.3.8.

This simulation study on the conjugate gradient as an efficient optimiser shows that all the regularisations introduced in section 3.1.4.3 are important to have reconstructions that better match the real object. The 3D reconstructions are not quantitative, but their 2D projections (OPD) are quantitative.

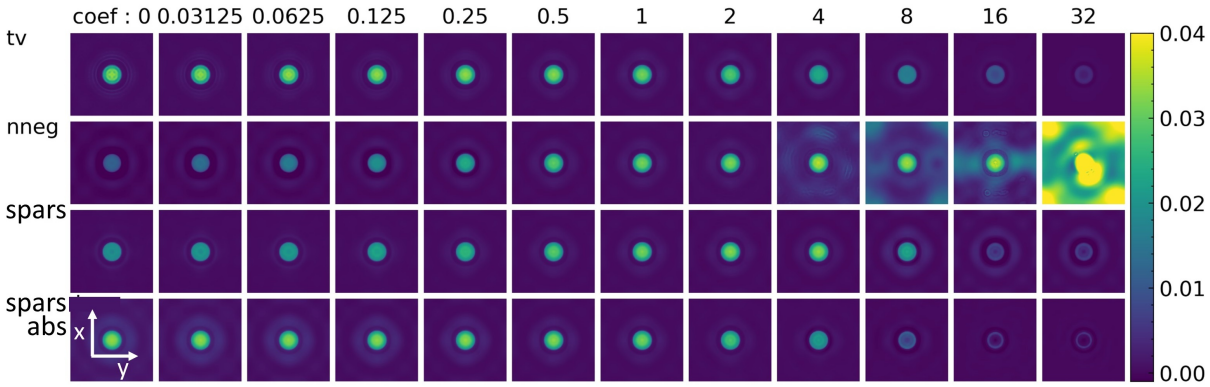


Figure 3.8: OPD of the reconstructed 3D bead from the simulated measurement

3.1.5.2 Iterative proximal Nesterov method

This new optimisation algorithm aims at minimising the same regularised inverse problem as defined in section 3.1.5.1:

$$\epsilon_{nneg,\beta,l1,l1_{abs}}(W) = \sum_{i=1}^{N_{img}} \left| |LightScat(\mu_{ix}, \mu_{iy}, W)| - \sqrt{I_i} \right|^2 + nneg.ReLU(-W_n) + \beta.TV(W) + l1.|W_n| + l1_{abs}.|W_{abs}|$$

The main difference is that this time the whole criteria will not be minimised as a whole, but it will be separated in 2 parts:

1. an accelerated gradient descent (with Nesterov) will handle the data fidelity term and the easy regularization terms for $l1$ and non-negativity.

$$\epsilon_{i,nneg,l1,l1_{abs}}(W) = \left| |LightScat(\mu_{ix}, \mu_{iy}, W)| - \sqrt{I_i} \right|^2 + nneg.ReLU(-W_n) + l1.|W_n| + l1_{abs}.|W_{abs}|$$

2. a proximal operator will be used to enforce the TV minimisation where a standard gradient is not very efficient. For total variation minimisation it is often shown that gradient are more unstable than proximal iterations [8].

This mixture optimisation algorithm is the following one:

Algorithm 6 Fast proximal gradient descent $(nneg, \beta, l1, l1_{abs}, \eta)$

Initialize $W_1^1 = 0$ and $b_0 = 1$ and $p_0 = p_1 = 0$
for $n_i = 1, 2, \ldots, N_{iters}$ **do**
  $b_{n_i} = \frac{1+\sqrt{1+4b_{n_i-1}^2}}{2}$
  $W_1^{n_i} = p_{n_i} + \frac{b_{n_i-1}-1}{b_{n_i}}(p_{n_i} - p_{n_i-1})$ – Nesterov acceleration
  for $i = 1, 2, \ldots, N_{img}$ **do**
    $\epsilon = \epsilon_{i,nneg,l1,l1_{abs}}(W_i^{n_i})$ – Compute the data fidelity error and the l1 and non-neg reg.
    $\Delta_i = \frac{\partial \epsilon}{\partial W_i^{n_i}}$ – Compute gradient for image i
    $W_{i+1}^{n_i} = W_i^{n_i} - \eta(\Delta_i)$
  end for
  $p_{n_i+1} = Prox_{\beta\|\cdot\|_{TV}}\left(W_{N_{img}+1}^{n_i}\right)$
end for

With $\text{Prox}_{\beta\|\cdot\|_{TV}}(x) = \arg\min_u \left\{\frac{1}{2}\|u-x\|^2 + \beta\|u\|_{TV}\right\}$

This algorithm 6 has been tested against the same optimisation problem as in the previous section 3.1.3: a 2 µm bead of refractive index 1.34 embedded in water with 84 intensity images with anlges corresponging to the data from our prototype.

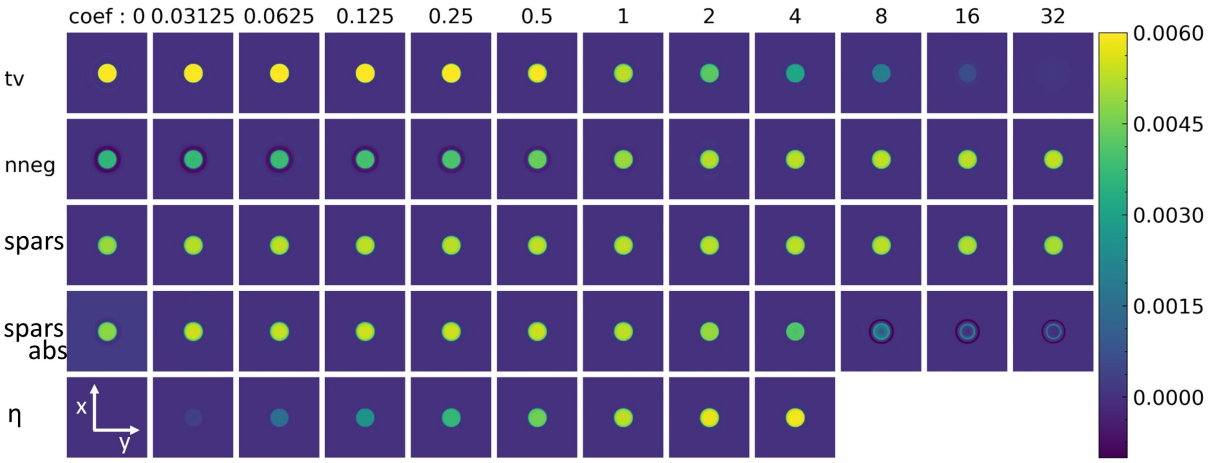


Figure 3.9: xy slice of the reconstructed 3D bead from the simulated measurement

Even though the optimisation problem is the same, with this approach the effect of each parameter is slightly different. This time, the initial values that produce good 3D reconstruction are slightly different (see below), mainly because the regularisation is performed

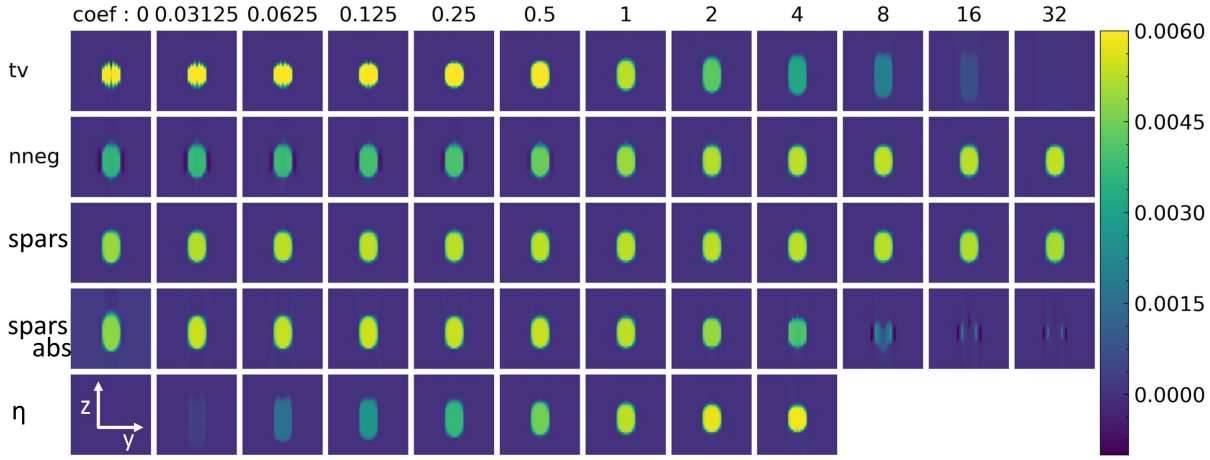


Figure 3.10: yz slice of the reconstructed 3D bead from the simulated measurement

at every angle. The TV regularisation inside the proximal is also different because it does not compete against the data fidelity but against a difference with the previous iteration object (the $\frac{1}{2}\|u - x\|^2$ in the proximal definition). The parameters used that gave satisfying results are: $\beta_{man} = 0.0004$, $nneg_{man} = 0.25$, $l1_{man} = 0.1$, $l1_{abs_{man}} = 2.5, \eta_{man} = 2$. To have a better understanding of the effect of each parameter **we will tune it again from zero up to 32 times the initial value.** The reconstructed object will change from not regularized enough to overly regularized.

The first and most important thing to observe is that the TV and η row are symmetrical. It was expected since the TV is working against the data fidelity term in alternations. Weakening the TV normalization will have the same effect as increasing the step of the optimisation algorithm. On the other hand, lowering the η will result in the same object as the one obtained after increasing the β. The main difference is that once the step is too high, the algorithm diverges and no results are produced (see the empty columns over the last row in Figures 3.9 and 3.10).

The TV effect on the object is sharper than previously, as it tends to fill the object with a constant value. A side effect of this discrete total-variation defined in equation 3.7 is that it tends to make the object wider and less intense in terms of RI. In the X and Y direction it is not an issue since the data fidelity imposes sharp refractive index gap. Along the Z

direction there is less information in the measurement to inform the algorithm on where the object should start or end. Since the TV norm tends to make the object as elongated as possible, as long as the data fidelity remains good, for high value of TV we have very elongated objects. This is an issue with our compact and low angular coverage prototype, because under low angle illumination most of the diffraction patterns comes from the OPD of the object and not from its shape along the Z dimension. An idea to reduce this elongation would be to change the L1 norm the gradient in each direction by a L0 norm. However, this is not technically feasible, as gradient descent cannot deal with L0 norm. A compromise would be to use a L0.9 or L0.8 norm, but it is already very unstable and quite challenging to use in a sens that the algorithm can diverge rapidly.

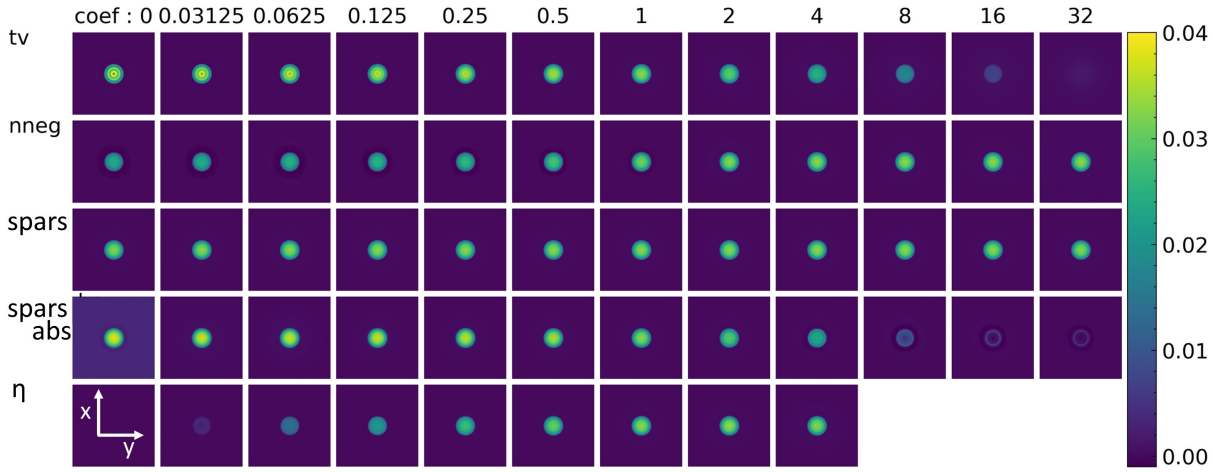


Figure 3.11: OPD of the reconstructed 3D bead from the simulated measurement

As it can be seen in Fig. 3.10 the non-negativity produced the desired result, i.e. it removes the negative voxel values in the reconstructed volume. This time the implementation is more stable and no undesirable effects (like strong hallucinated shape) happen even when the regularisation is 32 times stronger than the baseline.

Interestingly, the sparsity has almost no effect, the object looks sightly more elongated when it is removed (coef 0) but otherwise the choice of this parameter is not critic.

Finally the sparsity on the absorption seems to have a constant effect as long as it is not

pushed down to zero, where an elongation and a lowering of the refractive index happens.

Both accelerated gradient descent algorithms behaved much better than the naive gradient descent. Globally, the fast proximal gradient descent algorithm 6 provides better results than the conjugate gradient in terms of a quantitative OPD values Even objects without noise have an OPD close to the theoretical value of 0.04. The final reconstructions also show more robustness, as a wrong estimation of $nneg, \beta, l1, l1_{abs}$ has less influence on the reconstructed object. For example, the $nneg$ or the sparsity can be set factor of 30 higher without causing any significant issue on the reconstructed object (not the case for the conjugate gradient approach).

This sphere simulation emphasizes the crucial role of the regularisation hyper-parameters in the reconstruction and the necessity to choose them carefuly in the reconstruction process.

For stability reasons the Iterative proximal Nesterov method from section 3.1.5.2 will be used in the remaining of the manuscript for simulation and for biological sample.

3.2 TIMELAPSE PROTOTYPE

Our prototype conception was motivated by the will to study small developing organisms in 3D, from unicellular organisms to small animal models such as $Caenorhabditis\ elegans$ [10] as well as organoids and preimplantation embryos. This requires imaging in three dimensions on a broad range of length scales, from micrometer to study subcellular structures, to millimeter to analyse whole organisms. As detailed in the Introduction 1, we chose to use a non-fluorescent technique where we exploit a contrast due to light diffraction as in optical diffraction tomography (ODT [44] [88]) instead of fluorescence. Results of this section have been partially published in [99] and orally presented at [97].

3.2.1 Illumination and imaging component choice

To study living organisms one must be able to do so in the standard conditions where biological sample are grown. Most of these cultures take place in an incubator, a device maintaining specific temperature, humidity and CO_2 levels for growing and conserveing cell cultures as well as microbiological cultures. Incubators are therefore essential for the physiological development of the sample. That is why we made the choice to conceive a microscope able to work inside the incubator. The two main limitations are humidity and size: standard incubators are as big as a small fridge, leaving a maximum usable volume around 50×50×50 cm^3. It explains why our prototype has to be compact, and the smaller the better because incubators may be used for several different experiments at the same time and all the volume inside is not always accessible.

Another issue of long-term imaging of image biological samples is phototoxicity, especially for fragile samples like embryos. The amount of photons we target toward the sample must not interfere with the physiological development of the sample. Even though the question on how much phototoxicity affects the biological cultures is complex, some good practice emerged [71] and a proper quantification of this phenomenon should be performed before concluding any biological results. Considering the volume constraint as well as phototoxicity, a commercial LED array (SCI-Microscopy, US, $\lambda = 0.532\ \mu m$ for the green array, $\lambda = 0.625\ \mu m$ for the red array) was used for illumination. The phototoxicity can append with LED illumination, but the amount of photon that passes through the sample is much lower when using bright-field illumination instead of point scanning illumination where the whole sample has to probed like in confocal imaging.

Concerning the field of view, we aim to image biological samples that are thin enough for the collected light to be mainly ballistic and not diffused. Therefore we will limit us to biological object that are less than $200\ \mu m$ thick. Considering this constraint, two biological samples were selected to be imaged with our prototype, mouse embryos and liver organoids. Both have a spheroid shape, that is why we chose to limit our field of view to

$200\ \mu m \times 200\ \mu m$ in the xy direction. Considering a standard CMOS sensor with a $2.4\ \mu m$ pixel size and a 3088×2076 resolution (IDS UI3880), a 20X objective gives the desired field of view.

One last constraint that must be taken into account is a high sample versatility. A microscope should accept various shapes and thickness of sample containers. It excludes the use of immersion oil for simplicity of use, as well as the use of objective with high numerical aperture that usually have a short working distance (unless for very complex objective that comes at a premium price).

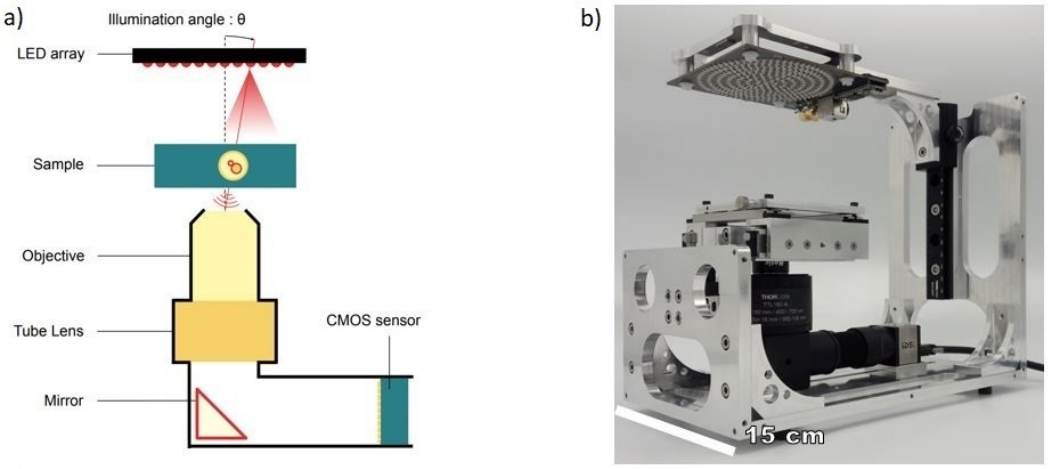


Figure 3.12: Overview of our compact L-shaped 3D microscopy setup

The final design of our prototype that fulfills all the previous requirements is presented in Fig.3.12. The illumination is achieved with a commercial LED array (SCI-Microscopy, US, $\lambda = 0.532\ \mu m$) centered at 84 mm above the objective. Only the first 133 LEDs are routinely used, providing a 0.346 NA illumination. With the goal to obtain a wide field of view ($\geq 100 \times 100\ \mu m^2$), we used a Motic $20\times$ 0.4 NA objective with a tube lens (Thorlabs, focal distance of 180 mm) to image on CMOS sensor with a 2.4 μm pixel size (IDS UI3880 - Schematic in Fig.3.12). This optical configuration results in a 0.120 μm effective pixel size. A 45° mirror has been added after the tube lens to have an "L-shaped" prototype that is more compact than an inverted in-line microscope.

3.2.2 Raw image normalisation

3.2.2.1 Integration time tuning

The CMOS sensor stores the gray value of each pixel in 8 bits, ranging from 0 to 255. To avoid saturation (value over 255 that will be decreased to 255) a good integration time must be chosen for each LED. The closer an LED is from the center of the LED array (that is aligned with the optical axis of the system), the brighter the image is. To characterise this effect we will study the mean value of an image with a constant exposition time of 30 ms.

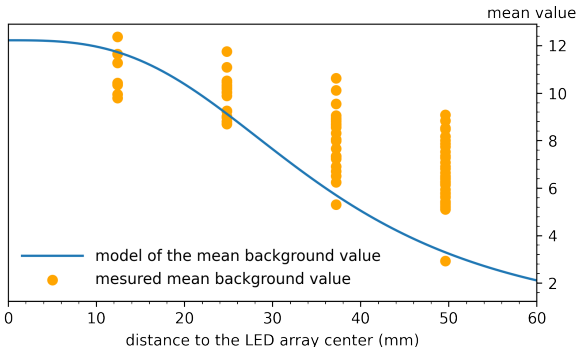


Figure 3.13: Evolution of the intensity of the light at the end of the optical system relatively to the distance to the center

On Fig.3.13 it is clear that a the mean pixel values decreases with the distance to the center of the LED array. Even more interestingly, the diversity of illumination intensity for a given distance to the center of the LED array is high. The brighter image for an LED that is placed at 50 mm from the center is 4 times higher than the darker one.

A simple model was fitted on the data with the goal to change the exposure to have a constant signal over noise ratio. Instead of a constant exposition time (et) of 30 ms, the exposition time was changed in function of the distance toward the center d by $et = 0.1 * (300 + \frac{d^3}{150})$

This adaptation to keep a constant signal to noise ratio on individual acquisitions and for any given illumination source changed the integration time between 80 and 800 ms,

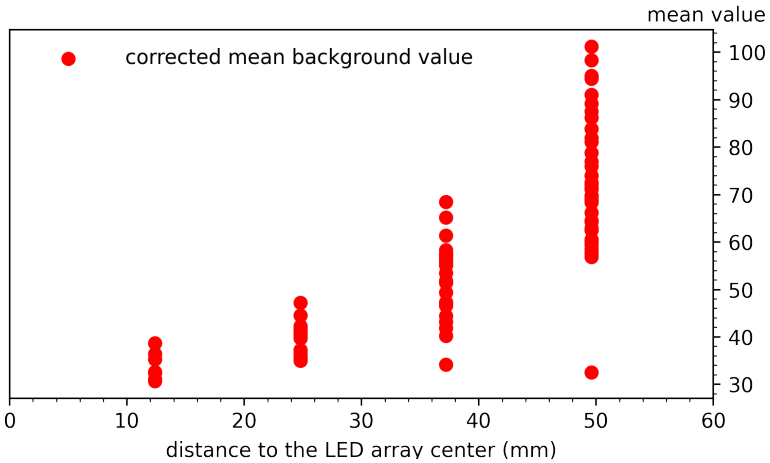
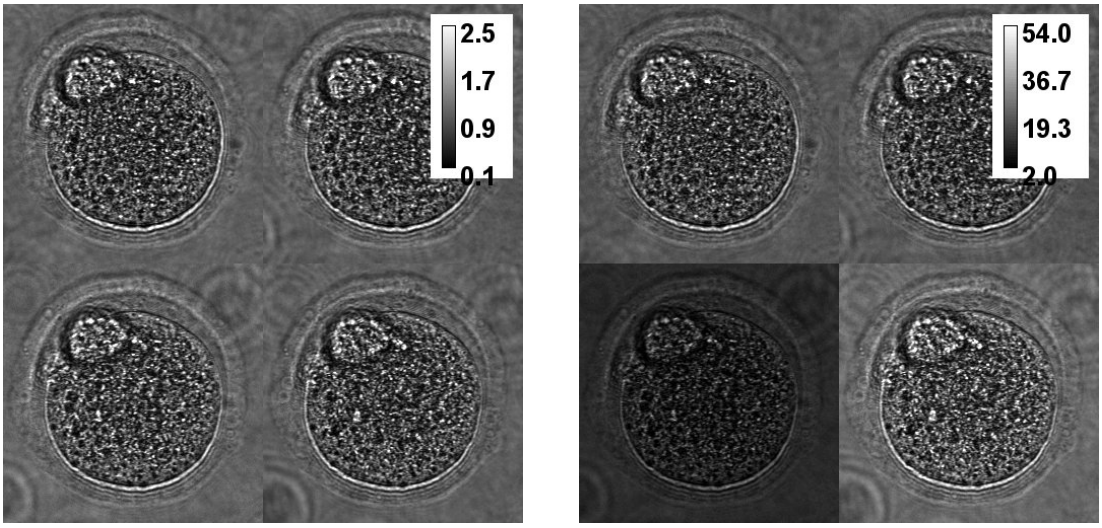


Figure 3.14: Intensity of the light at the end of the optical system relatively to the distance to the center with a corrected exposure time.

producing a total acquisition times of about 2 minutes for 133 frames.

3.2.2.2 *Prepossessing of the raw images*



(a) Numerically normalized **(b)** Un-normalized

Figure 3.15: Brightfield images of an embryo acquired with four consecutive illumination LED Fig.a) shows the images after normalization Fig.b) shows the raw images

For the reconstructions, an intensity background level close to unit is essential to match

the value of the incoming plane wave of our forward light scattering model (see Section 3.1.1). This means that if no object diffracts light, the acquired image should be a constant image with value one.

Even though the adaptation of the exposure times was a first step toward having normalised raw data, the variability inside each ring is too high and each image has to be treated individually. To have a constant background with the value 1, each image is divided by the mean value calculated over the overall intensity distribution of a manually selected crop of the image where no object is present. This guaranties that every image is correctly normalized, but requires a manual input from the user on the reconstruction algorithm.

3.2.3 Reconstruction on calibrated micro sphere

To validate our normalization procedure and confirm the first results obtain after simulations in section 3.1.5.2, we will first apply our reconstruction algorithms 6 to the reconstruction of calibrated micro-sphere. We chose 1 µm silica beads that have a refractive indice (RI) of 1.4570 at 625 nm. They are placed inside water with a refractive index of 1.33. It creates an OPD of 0.127 µm, slightly less than 2 times the wavelength. This is already considered as an optically thick sample, challenging enough for a first validation of the prototype.

To perform the reconstruction of these silica micro-spheres, 133 images were recorded with illumination angles ranging from 0° to 20.3° at 625 nm and a 20X/0.4NA objective. The Fast proximal gradient descent algorithm with 40 iterations and regularisation parameter $\beta_{man} = 0.0016$, $nneg_{man} = 2$, $l1_{man} = 0.15$, $l1_{abs_{man}} = 1.5, \eta_{man} = 8$ was used to produce results in Fig.3.16-e-f-g-h. On the z cross-sectioning of Fig.3.16-e the reconstructed bead is elongated up to 4 times its length along the z direction. This elongation is mainly caused by a limited angle coverage and a rather low numerical aperture of our objective compared to high NA objective (NA>1) as used in reference [143].

This elongation leads to a refractive index that is lower than expected inside the bead

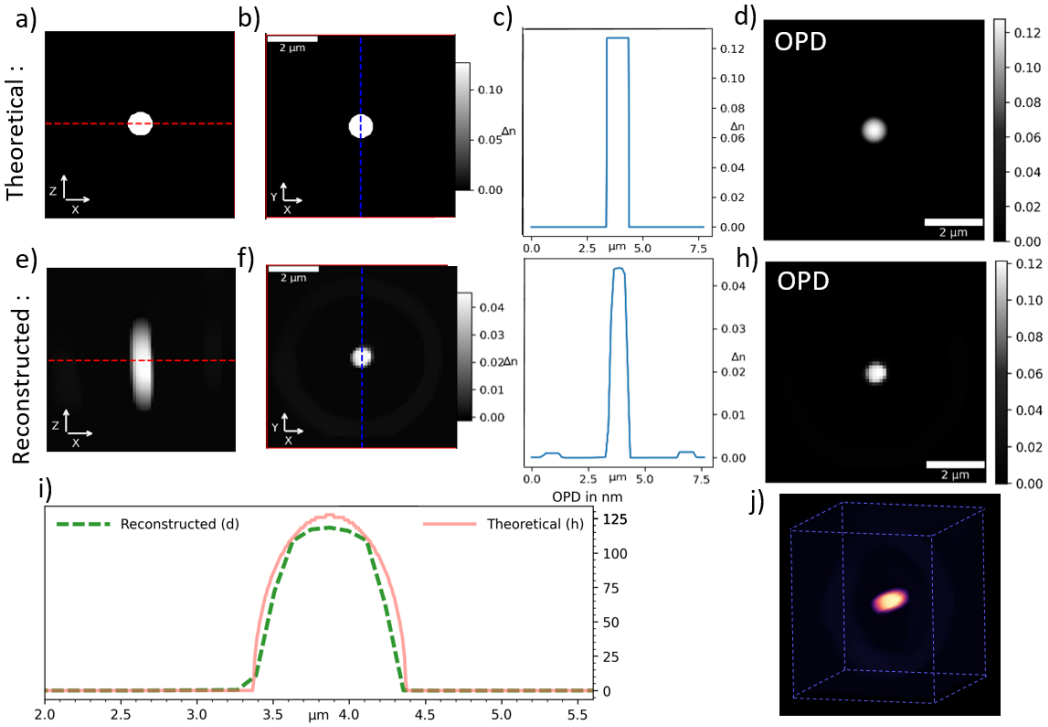


Figure 3.16: Reconstruction of a calibrated micro-sphere. a-b-c-d - Theoretical micro-sphere. e-f-g-h - Reconstructed micro-sphere. d, h, optical path difference (OPD) of the theoretical and reconstructed micro-sphere, respectively. i- Line plot of the OPD shown in Fig 3.16 d,h. j is a 3D rendering of the reconstructed micro-sphere.

(0.045 in Fig.3.16-g instead of 0.12 in Fig.3.16-c). At present our system does not yield quantitative results for every voxel reconstructed. Nevertheless if we consider the object as a whole, we can compute the optical path difference (OPD) on-axis (Fig.3.16-d,h). When comparing the micro-sphere reconstructed OPD with a theoretical value, we found on Fig.3.16-i a good agreement between the theory and our system.

Our final goal is to achieve time-lapse imaging of thick biological samples, therefore we cannot limit our study to calibrated object like micro spheres. The following sections will therefore describe the assessment of our imaging capability on biological living samples.

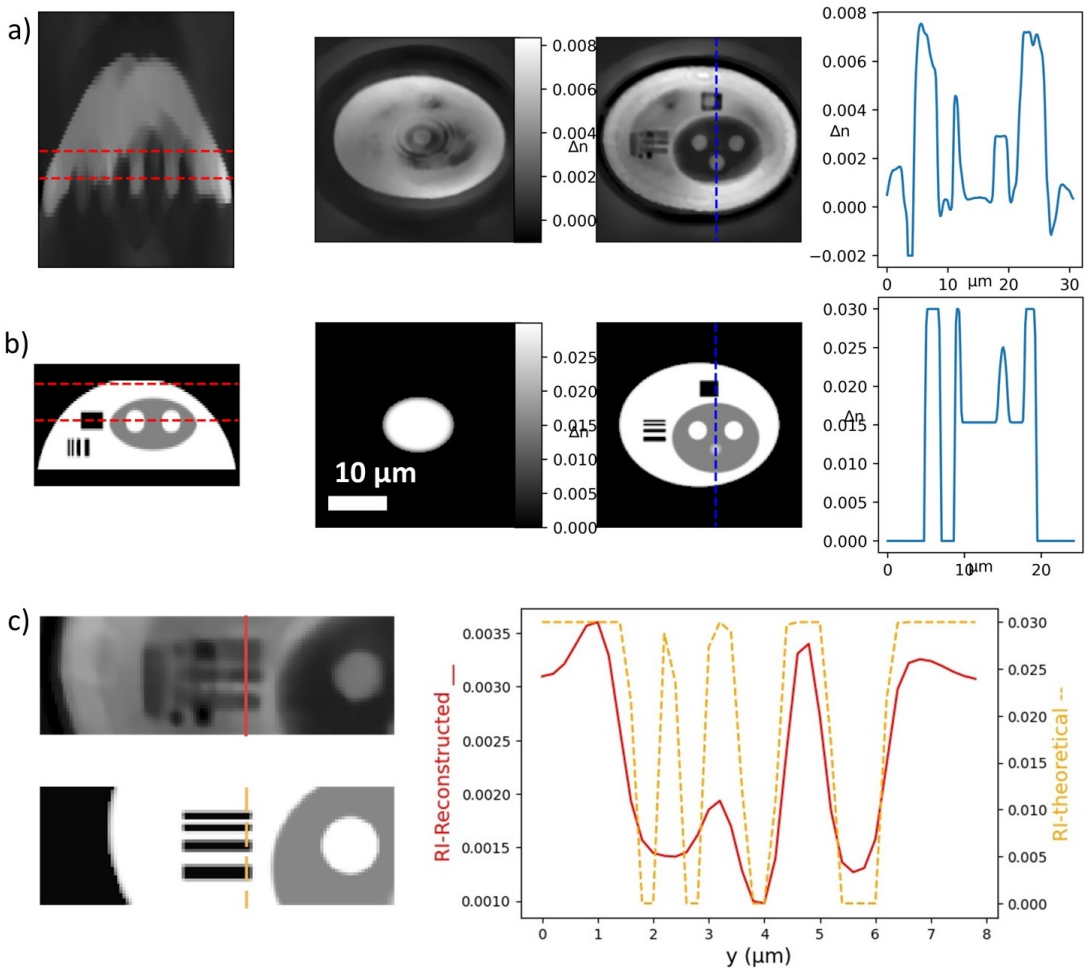


Figure 3.17: Resolution and shape characterisation on a 3D printed cell phantom. a- Shows the 3D reconstruction of the phantom. First image is the central zx slice. Second image is the yx section on the bottom of the phantom (position shown by the top dotted line on the zx slice). Third image is another yx slice (lower red dotted line). Finally the dotted blue line is plotted to show the refractive index distribution over a given section. b- Shows the 3D theoretical phantom. First image is the central zx slice. Second image is the yx in top of the phantom, position showed by the upper dotted line on the zx slice. Thirds image is an other yx slice (lower red dotted line). Finally the dotted blue line is plotted. c- Zoom on the Y resolution target, with the plot of the RI distribution

3.2.4 Lateral and axial Resolution

In microscopy, the notion of resolution is crucial and everyone expects a value for the xy direction, as well as the Z direction. For 3D optical diffraction tomography this question

does not always have a simple answer, like in any regular microscopic system. Ziemczonok, M., Kuś, A., Wasylczyk, P. $etal.$ proposed a new way to quantify phase imaging systems with novel 3D printed biological cell phantoms [137]. This technique it is based on photo-lithography, where the absorption of photons leads to the curing of a liquid polymer. This creates phase only 3D structures with a $100\ nm$ lateral and $500\ nm$ axial resolutions and given refractive index. This phantom was used to characterize the resolution of our setup (as detailed below), as a part of an active collaboration with the group in Warsaw University of Technology.

The phantom is a half dome of 30 $\times$ 20 µm on its base and 20 µm thick as shown in Fig.3.17 b). The main features of the phantom are the three spherical nucleoli placed in a region of lower (gray) RI that mimics a nucleus, which is often present in many cells. The features are here to access our ability to reconstruct shapes of large sub-cellular structures and their RI. Outside of this fake nuclei there are three resolution test targets with a strong RI gap along all the axis. These resolution target enable us to quantify the capacity of the system to reconstruct high RI gradients features, such as lipid droplets.

3.2.4.1 Lateral resolution

A total of 84 intensity fields were acquired with our prototype equipped with a $\times 20$ objective (NA = 0.40) and green LEDs with $\lambda = 530\ nm$. The reconstructions of this single cell that is $20\ \mu m \times 30\ \mu m$ and $15\ \mu m$ thick was performed with the Fast proximal gradient descent algorithm detailed in section 3.1.5.2.

As excepted with our low NA, the reconstructed image after intensity only acquisitions is strongly elongated along the Z direction. On Fig.3.17 a-, the first yx slice looks like it is in the middle of the reconstructed phantom, but we are truly on the top of the phantom because we can observe the top layer of the 3D printed top. This is not visible on the edges because the gap between each printed slice are really small, but on top, the distance between each printed slice is bigger and can be observed. Yet, the major key morphological characteristics of this phantom are correctly retrieved.

To answer the resolution question, we can start by defining exactly what is the resolution. Resolution is the minimum distance at which two distinct objects of a sample can still be seen as separate entities. The resolution target on Fig.3.17 c- is perfect to measure resolution. As shown in the image and in the RI distribution graph, the first two $300\,nm$ marks are indistinguishable and appear blurred as only one object. Then the $500\ nm$ hole can be separated sharply from the $700\ nm$ hole, while the graph show a slight blend over the $300\ nm$ hole. From this study we can affirm that our system has a **1 µm lateral resolution.**

The theoretical value from Fig.1.1 with our system composed of a 0.4 NA objective and $\lambda = 0.532$ is around $0.8\ \mu m$.

3.2.4.2 Axial resolution

In the previous section, we used the following definition for the resolution: "the **unique** minimum distance at which two distinct objects of a sample can still be seen as separate entities". Since our major issue is the axial elongation, and that this elongation has the shape of cone, the width of the object has a huge impact on how far will this cone propagate.

For example if we stack two $2\ \mu m$ beads from section 3.2.3, they would appear as separate entities if we place them at $6\ \mu m$ from each other. But if we use smaller object, like a $1\ \mu m$ sphere, we could bring them closer before their elongation cones hit each other, $3\ \mu m$ in this case. If we consider bigger objects like embryos that are $100\ \mu m$ wide, they need to be placed more that $300\ \mu m$ away so that it can be seen as separate entities.

The question remains, what is the axial resolution of our system ? $3\,\mu m$? $6\,\mu m$? $300\,\mu m$?

3.2.5 Fixed embryo 3D reconstruction and aberration correction

Most of the cell cultures are essentially 2D, the cells are grown on a slice or any other cell container and remain mainly flat as they develop. In order to access more complex sample

a collaboration with the hospital in Grenoble[1] was initiated, allowing us to image mouse embryos. This is the perfect sample to validate our microscope with, since embryos vary between 100 and 150 μm with many 3D complex internal structure. The team at Grenoble hospital was responsible for all embryo collection and manipulation.

3.2.5.1 Embryo preparation and transportation

Embryo preparation

Zygotes were collected from mature C57BL/6 females, synchronized with 7.5 units of pregnant mare serum gonadotrophin (PMSG, MSD santé animale # Chronogest 600) and 7.5 units of human chorionic gonadotrophin (hCG, MSD santé animale # Chorulon 1500), mated with fertile males. Oviducts were removed in M2 medium (Sigma-Aldrich # M7176) and zygotes were collected from the ampulla. Cumulus cells were removed with hyaluronidase 0.1 mg/ml. Embryos were then washed and maintained until blastocyst stage in KSOM medium (Merk MR-106-D) in an incubator at 5% CO_2 and 37°C. The incubator used is a SANYO CO2 incubator MCO-19AIC.

Animals ethics

All animal procedures were run according to the French guidelines on the use of animals in scientific investigations with the approval of the local Ethics Committee (ComEth Grenoble N° 318, ministry agreement number #7128 UHTA-U1209-CA). All animals (C57BL/6) were from Charles River laboratories. All animals used were 5-8 weeks old for females and 2-6 months old for males.

Embryo transportation

During the first month of test, the embryos were prepared in the hospital by the biologist and transported in our lab to fine tune the microscope. Since it was not allowed to bring living organism cultivated in another lab, we had to fix those embryo with a solution of paraformaldehyde. This makes the embryos more stable over time, they can be easily

[1]Univ. Grenoble Alpes, INSERM U1209, CNRS UMR 5309, Institute for Advanced Biosciences, Team Genetics Epigenetics and Therapies of Infertility, 38000 Grenoble, France.

transported without the need to have a controlled environment (37°C and 5%CO_2) but their true shape and all the inner structures are unmodified. They still need to be maintained in a KSOM medium (Merk MR-106-D) not to dry out. Routinely they are stored in a Lab-Tek® container with $0.5\,mm$ on oil on top of the solution to prevent evaporation. This kind of container must be transported carefully to prevent the medium or the oil to spill, and it was not possible to guarantee it while crossing the city from the hospital to our lab.

To guarantee that the sample holder would not spill, we chose to transport the fixed embryos in $20\,\mu m$ LEJA slide sealed with synthetic resin. These slide have a $1000\,\mu m$ glass thickness on one side and $720\,\mu m$ on the other side. It produced a stable sample that had almost no changes in its shape of internal structure up to a week, that could be transported easily without particular care and that was fixed (i.e. not living anymore).

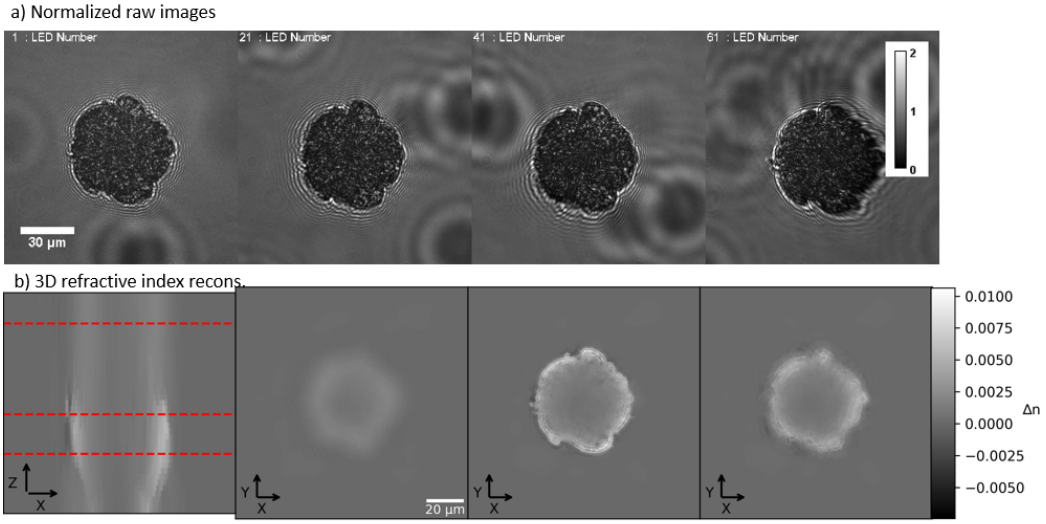


Figure 3.18: Reconstruction of a multi-cell embryo trough a thick ($720\,\mu m$) glass sample a- The raw images acquired with our microscope for various LED. b- The 3D refractive index reconstruction with 3 different xy slice. Each slice is positioned with a dotted red line on the zx slice (from top to bottom, respectively).

Even though the transport was robust, the first images with this sample preparation were disastrous. The raw images in Fig.3.18 a show a speckle-like pattern for every imaging angle, and only the contour of the object looks like a standard diffraction pattern with

fringes. After reconstruction (Fig.3.18 b), the center of the embryo is almost constant with a zero value, meaning that no phase information could be retrieved from the raw images. As expected, only the edges are sharp and visible.

3.2.5.2 *Aberration correction - Theoretical formulation*

It took me a few months to understand why we had the disastrous results from the previous section 3.2.5.1. What was causing those aberrations is the glass thickness between the object and the objective. Normally, standard objective are design to work with standard coverslip that are 170 μm thick. In our case the LEJA slide was 720 μm thick.

A theoretical study on the effects of the glass thickness was thus conducted to have a better idea of the effect and strength of the aberrations.

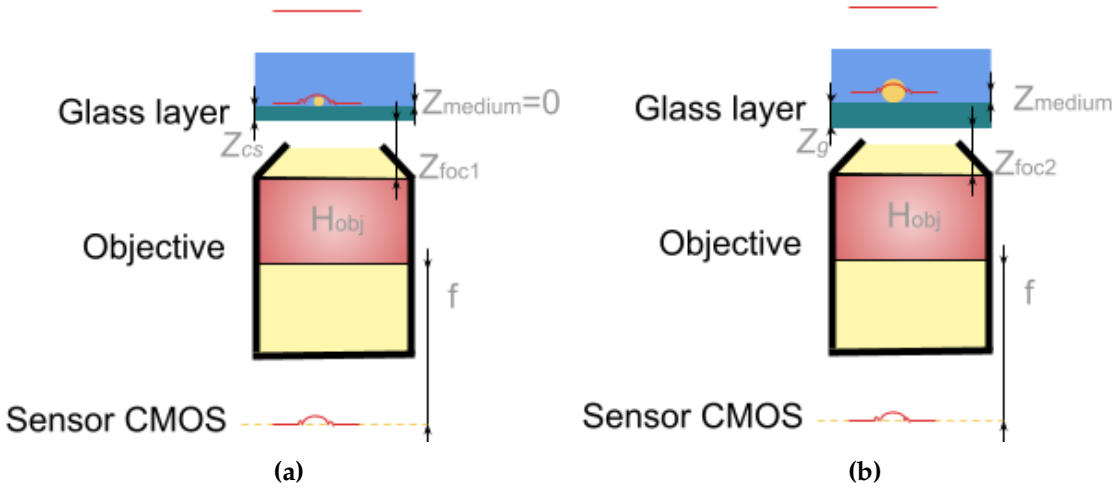


Figure 3.19: Schematic view of a microscope interface system. a) Representation of the nominal configuration of the objective. b) Representation of the configuration with a custom sample container and thick object.

If we consider the objective in its nominal utilisation, it is supposed to accurately reproduce the object to image on the CMOS sensor (except for the high frequency that are cut by the numerical aperture). This can be expressed with the following equation:

$$1 = H_{obj}(\vec{\mu}_{2D}) e^{2i\pi(z_{cs}\sqrt{(\frac{n_g}{\lambda})^2-\mu_{2D}^2}+(z_{foc}+f)(\sqrt{(\frac{n_a}{\lambda})^2-\mu_{2D}^2})} \tag{3.11}$$

From the previous equation we can extract an analytical formulation of the full objective transfer function and inject it inside the general case equation:

$$H_{global}(\vec{\mu}_{2D}) = p(\vec{\mu}_{2D})e^{2i\pi(z_m\sqrt{(\frac{n_m}{\lambda})^2-\mu_{2D}^2}+(z_g-z_{cs})\sqrt{(\frac{n_g}{\lambda})^2-\mu_{2D}^2}+(z_{foc2}-z_{foc1})(\sqrt{(\frac{n_a}{\lambda})^2-\mu_{2D}^2})} \tag{3.12}$$

The $z_{foc2} - z_{foc1}$ remains to be expressed. In both situation (a) and b) from Fig.3.19) the field propagates equally in the Fourier space, this means that the propagation before entering the objective have to be equal between a) and b):

$$\begin{aligned} &e^{z_m(2i\pi\sqrt{(\frac{n_m}{\lambda})^2-\mu_x^2-\mu_y^2})}.e^{z_g(2i\pi\sqrt{(\frac{n_g}{\lambda})^2-\mu_x^2-\mu_y^2})}.e^{z_{foc1}(2i\pi\sqrt{(\frac{n_a}{\lambda})^2-\mu_x^2-\mu_y^2})} \\ &= e^{z_{cs}(2i\pi\sqrt{(\frac{n_g}{\lambda})^2-\mu_x^2-\mu_y^2})}.e^{z_{foc2}(2i\pi\sqrt{(\frac{n_a}{\lambda})^2-\mu_x^2-\mu_y^2})} \end{aligned} \tag{3.13}$$

By using the notation $\mu_\lambda = \lambda^2(\mu_x^2 + \mu_y^2)$ we can simplify the previous equation to:

$$z_m n_m\sqrt{1-\frac{\mu_\lambda}{n_m^2}} + z_g n_g\sqrt{1-\frac{\mu_\lambda}{n_g^2}} + z_{foc1} n_a\sqrt{1-\frac{\mu_\lambda}{n_a^2}} = z_{cs} n_m\sqrt{1-\frac{\mu_\lambda}{n_g^2}} + z_{foc2} n_a\sqrt{1-\frac{\mu_\lambda}{n_a^2}} \tag{3.14}$$

After noticing that $\mu_\lambda << 1$ every $\sqrt{1-x}$ can be approximated by $1 - \frac{x}{2}$ as long as $x << 1$:

$$\frac{z_m}{n_m} + \frac{z_g}{n_g} + \frac{z_{foc1}}{n_a} = \frac{z_{cs}}{n_m} + \frac{z_{foc2}}{n_a} \tag{3.15}$$

Then the $z_{foc2} - z_{foc1}$ distance expression is straightforward:

$$z_{foc2} - z_{foc1} = n_a(\frac{z_m}{n_m} + \frac{z_g - z_{cs}}{n_g}) \tag{3.16}$$

by replacing it inside equation 3.12 we get:

$$H_{global}(\vec{\mu}) = p(\vec{\mu})e^{2i\pi(z_m\sqrt{(\frac{n_m}{\lambda})^2-\mu^2}+(z_g-z_{cs})\sqrt{(\frac{n_g}{\lambda})^2-\mu^2}-n_a(\frac{z_m}{n_m}+\frac{(z_g-z_{cs})}{n_g})(\sqrt{(\frac{n_a}{\lambda})^2-\mu^2})} \tag{3.17}$$

This new transfer function 3.17 gives us the possibility to account for a non-standard coverslip thickness as well as a various depth inside the container, making our reconstruction algorithms quite versatile and adjustable to different applications.

3.2.5.3 *Aberration correction - Effect on calibrated micro sphere*

To make sure that this new transfer function mimics correctly the way a glass layer distorts the electromagnetic field, we imaged $2\ \mu m$ silica beads with given optical index $n = 1.46$ (at $\lambda = 532\ nm$) in water. The solution was placed in a custom container that produced 2 different aberrations on each side. On one side the interface thickness is $170\ \mu m$, one the opposite side it is $1000\ \mu m$. This was a custom build with a $170\ \mu m$ coverslip taped on top of a standard microscope slide. The comparison between the light scattering model and the true acquisition from the prototype are detailed in Fig.3.20

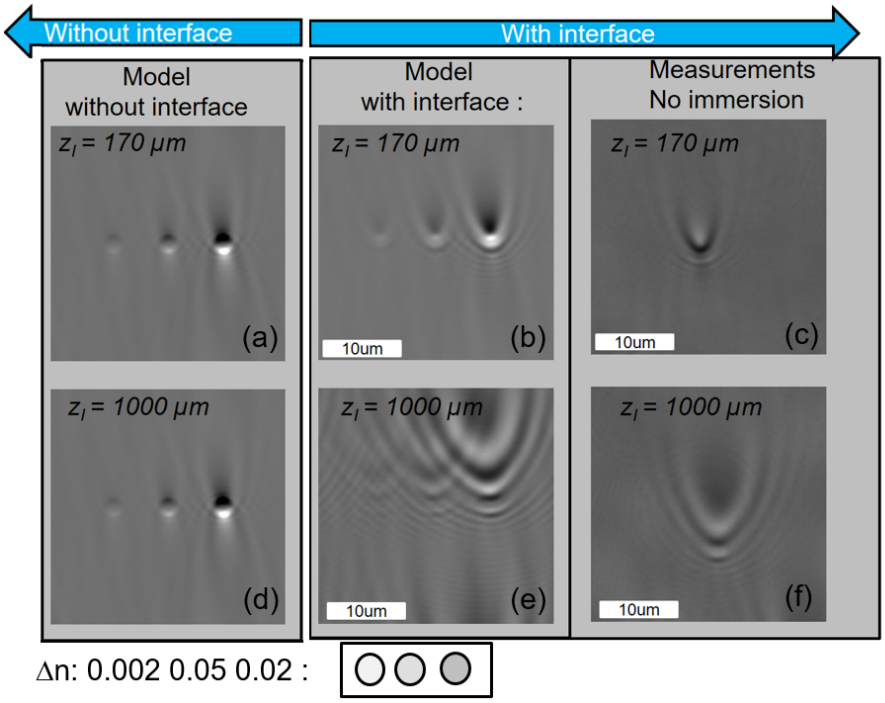


Figure 3.20: Comparison between 2 µm silica beads of given n obtained with simulations at various thicknesses of interface and with microscopy acquisitions using an illumination angle of 15° in air. a-b-d-e are results of the simulation with 3 spheres of refractive index gap 0.002, 0.05 ,0.02 (from left to right) c-f are the real image from our microscope. Z_i denotes the interface thickness

As expected, Fig.3.20 a-d are identical, since this model does not account for interface aberration. When the interface is taken into account, it is clear that the shape of the

diffracted wave after the silica bead better matches our microscopy acquisitions. Indeed, Fig.3.20 e has thin white and black fringes bellow the bead, while the fringes on the top are wider and more distant from each other, similarly to the brightfield image from the microscope in Fig.3.20 f.

It appears that our multiple scattering model is more precise when we take into account the aberrations introduced by the non-standard interfaces. However, this does not automatically guarantee that the reconstructions will be improved as well. To prove this point, we decided to conduct a study on a phase calibration slide.

3.2.5.4 Aberration correction - Experiment on calibrated phase USAF targets

We emphasized that in the case of non-standard coverslip thicknesses or sample depths inside the imaging medium it is important to take into account the spherical aberrations in the direct model. This will be further shown herebelow on measurements performed with a $1000\ \mu m$ thick USAF phase calibration slide.

On a $1000\ \mu m$ thick microscope slide, $300\ nm$ of glass have been etched everywhere except on the USAF target patterns. It creates a really thin, characterized, phase target with an Optical Path Difference (OPD) of $150\ nm$.

This calibration sample has been imaged in the same condition as the calibrated microsphere. 133 images were recorded with illumination angles ranging from 0° to 20.3° at $625\ nm$ and a 20X/0.4NA objective. At $625\ nm$ the glass has a refractive index (RI) of 1.5154. The glass is placed in the air creating a refractive index gap of 0.5154 on $300\ nm$, resulting in a theoretical optical path difference of $154\ nm$. The fast proximal gradient descent algorithm 6 with 40 iterations was used to produce results in the Fig.3.21-a-b.a

The reference measurements have been acquired with a quadriwave lateral shearing interferometry (QLSI) integrated in a commercial Phasics camera. Measurements were performed on a conventional inverted microscope (Zeiss Observer Z1) with 40× 0.64 NA air objective using Kohler illumination with a $750\ nm$ long pass filter (AHF F32-750E). A wavefront sensor SID4Bio (Phasics, Saint-Aubin, France) was mounted on the video port

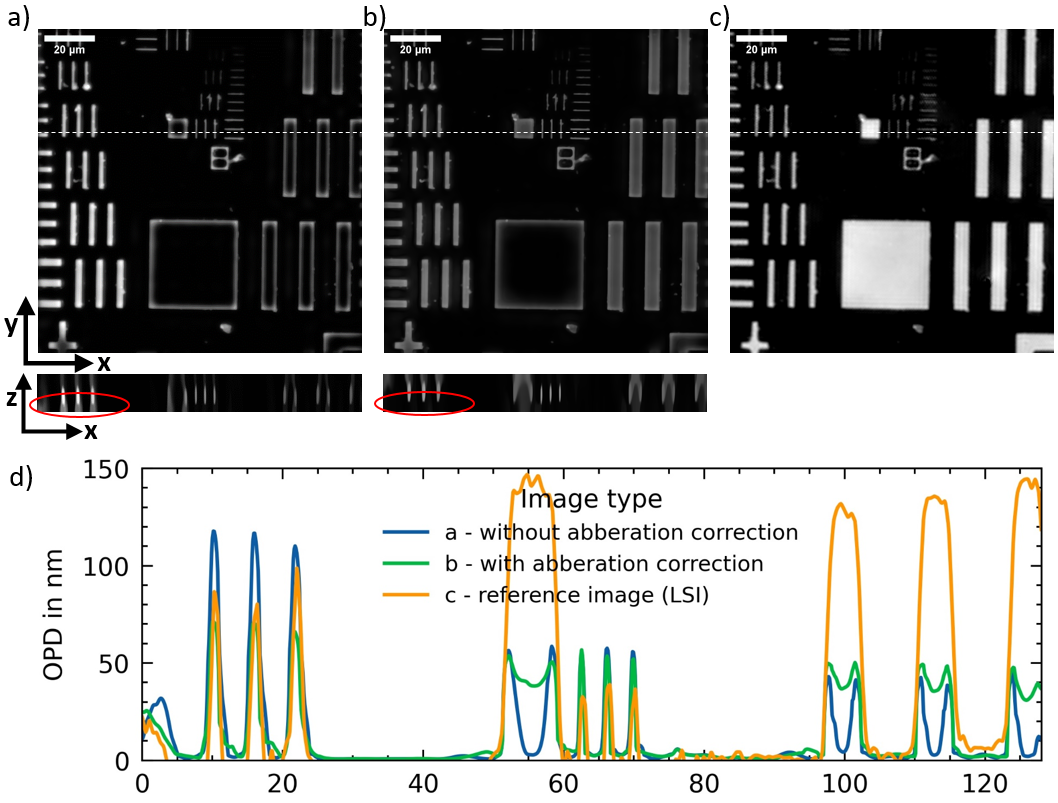


Figure 3.21: Reconstruction of a USAF phase target. a- Reconstruction without using the aberration kernel. b- Reconstruction with the aberration kernel. c- Reference image (QLSI). d- OPD plot of the dotted line in Fig. a-b-c. Scale bar is 20 µm wide.

of the microscope. The wavefront sensor SID4Bio contains a 2D grating (modified Hartmann mask) placed in front of the camera. The grating replicates the incident wavefront and after a short propagation an interferogram is recorded on a CCD camera. The interferogram is analyzed in real time by Fourier transforms to extract intensity in 0 order and OPD gradients in 1 orders, along X and Y directions. The latter are finally integrated in two dimensions to yield wavefront measurement, which is the OPD in the projective approximation.

Intensity only measurements only contain information about the variation of the refractive index. This is why the high spatial frequencies of objects are easier to reconstruct than the lower spatial frequencies. This can be observed on the big square in Fig.3.21 a and b, where only the edges are visible. To recover the lower spatial frequencies, we need

an accurate light scattering model that can perceive differences between an "edge only" object and a "full object", in addition to a regularization that favors continuous and "full object" like the total variation (TV) norm. When the spherical aberrations are not taken into account, the light scattering model is not precise enough and the TV norm cannot fill the rectangle object on the right of the image in Fig.3.21 a. When spherical aberrations are taken into account, the rectangles on the right in Fig.3.21 b are more filled, even though the edges are still higher that the center as plotted on Fig.3.21 d at $100\,\mu m$ and $115\,\mu m$ with the green curve.

When looking at the zx slice in Fig.3.21 a-b we can observe that the elongation along the z direction is less pronounced (red ellipse) when using our aberration kernel.

This USAF phase target is far from biological objects for which our microscope has been designed. Indeed when we have a biological sample with an OPD of $150\ nm$, it often measures $15\ \mu m$ with a refractive index of 1.34 inside a water of refractive index 1.33. This creates a refractive index gap of 0.01 and the fact that this gap is much lower that the refractive index of the medium (here water at 1.33) is part of the approximations of our light scattering model. In this USAF phase target the refractive index gap is 0.5, this might be one of the reasons why our OPD measurements are 3 times lower than the reference measurements. Even though we are not phase-quantitative on this phase target, the morphology was correctly retrieved.

The last step to validate the aberration correction algorithms is to show their relevance when applied to biological sample.

3.2.5.5 Aberration correction - Reconstruction on thick biological objects

After validation on a calibration slide we tackled a more challenging sample, a multi-scattering 5 days embryo. In Fig.3.22, its cross-sections are shown, with and without taking into account spherical aberrations. The embryo was placed inside a LEJA slide (Ref 026857 from IMV Technologies - LEJA) with a thickness of 720 μm (by thickness we mean the Z_g distance defined in Fig.3.19). The non-standard glass thickness created spherical

aberrations that are taken into account with the kernel detailed in section 3.2.5.2.

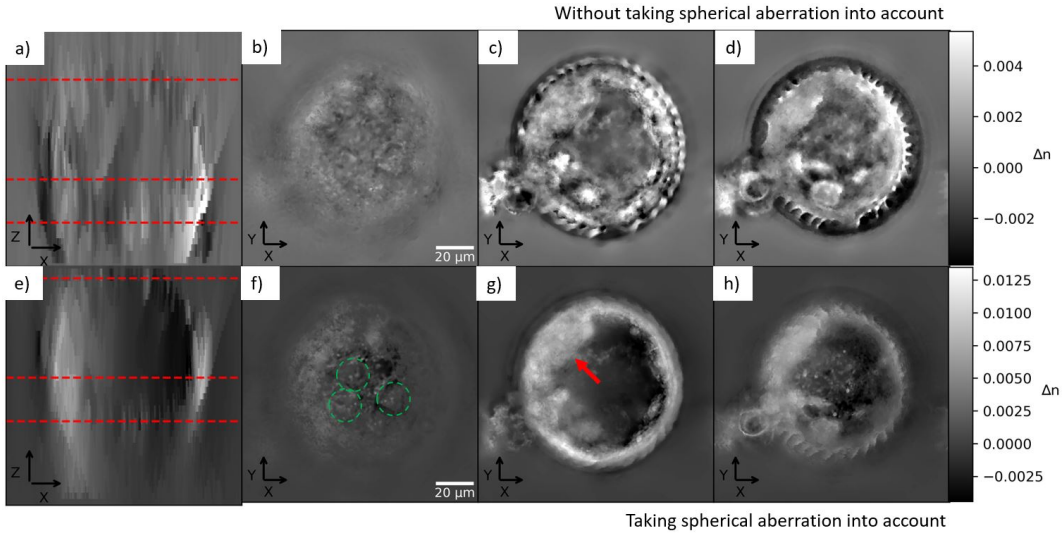


Figure 3.22: Reconstruction of a thick, multiple scattering embryo at blastocyst stage. e,a - Axial view of the reconstructed embryo with and without taking into account spherical aberrations, respectively. b,f - cross section view at $z = -28.35\ \mu m$. c,g - cross section view at $z = 9.45\ \mu m$. d,h - cross section view at $z = -31.5\ \mu m$.

Fig.3.22g shows a blastocyst composed of more than a hundred cells, mostly found in the inner cell mass in the top left (red arrow). This creates a cavity (the blastocoel) that is almost empty, which can only be distinguished in Fig.3.22g. Taking into account the aberrations allows to identify the typical trophoblast cells, composing the periphery of nominal developing blastocysts (green dashed circles one on Fig.3.22 f).

The external structures are correctly retrieved with our device without correction of spherical aberrations for the top part of this embryo (see Fig.3.22 b). However, as we move deeper inside the embryo (toward negative z, Fig.3.22 d) the optical aberrations produce artifacts that are not physiologically present in healthy embryos. This is highlighted by the comparison between Fig.3.22 c and 3.22 g or between Fig.3.22 d and 3.22 h.

3.2.6 3D time lapse of *in-vivo* embryo

The set up was further used in real-time directly in a cell culture incubator to unveil embryonic development in physiological conditions. Even if the lower illumination angle

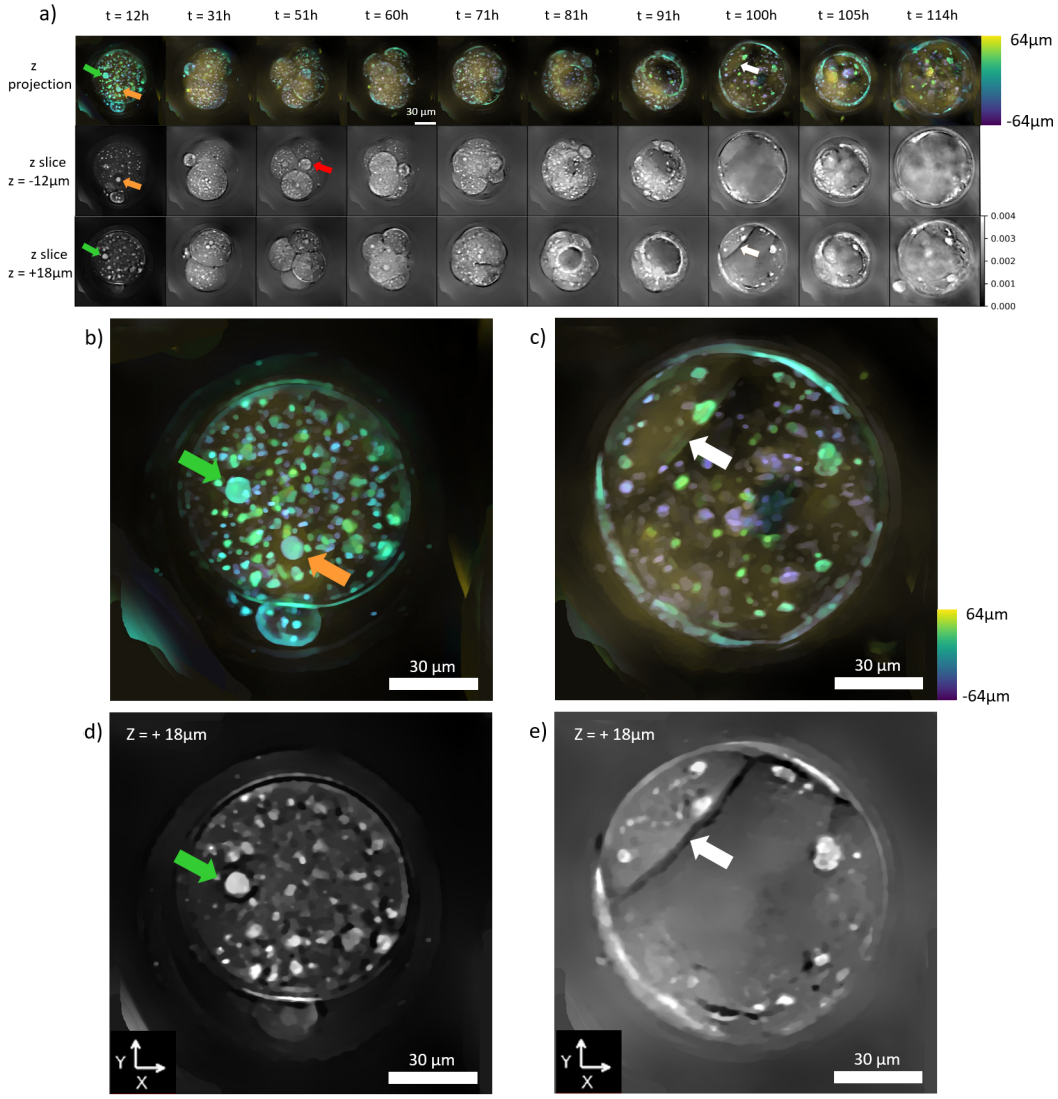


Figure 3.23: Time lapse images of a mouse embryo developing from fertilization to expanded blastocyst a) Illustrations of images taken at different developmental time and corresponding to different analyses (3D, Z projection, $z = -12\ \mu m$ and $z = +18\ \mu m$). The color code on figure one the z projection represents the position along the z axis. See Visualization 1 for all the frames. Scale bar is $30\ \mu m$. b) is the z projection a t = 12 h. c) is the z projection at t = 100 h d) is a z slice at t = 12 h. e) is a z slice at t = 100 h. Orange and green arrow point the two pronuclei that are in different z positions, the red arrow points the polar body and the white arrow points the inner cell mass of the blastocyst.

and the simple optics degrades the resolution, it brings new possibilities in real-time monitoring. Embryos were prepared as described in Section 3.2.5.1 and seeded in a 16-dish

EmbryoSlide (Vitrolife). The acquisition was performed with 84 illumination angles every hour for 6 days on a growing mouse embryo and the reconstructions were performed with the fast proximal gradient descent algorithm detailed in section 3.1.5.2 with 40 iterations. The reconstructed volume was $1024 \times 1024 \times 128$ big, resulting in a 35 minutes reconstruction time for each volume on a non-professional GPU, a NVIDIA GEFORCE RTX 3090.

Fig.3.23 and visualization 1 show all the major steps of embryo development, from an initial stage of one cell zygote to a blastocyst. Even-though we cannot truly know when fertilization occurred, we will consider that the first mitosis happens 20 h after the fertilization and all the timing in this section will take the seeding as t = 0. At t = 12 h, we can identify the 2 pro-nuclei that will merge during embryo development. The male and the female pro-nuclei are not in the same z position, the one pointed with the orange arrow can be seen in the $-12\ \mu m$ optical slice where the one pointed by the green arrow is more located around $+18\ \mu m$.

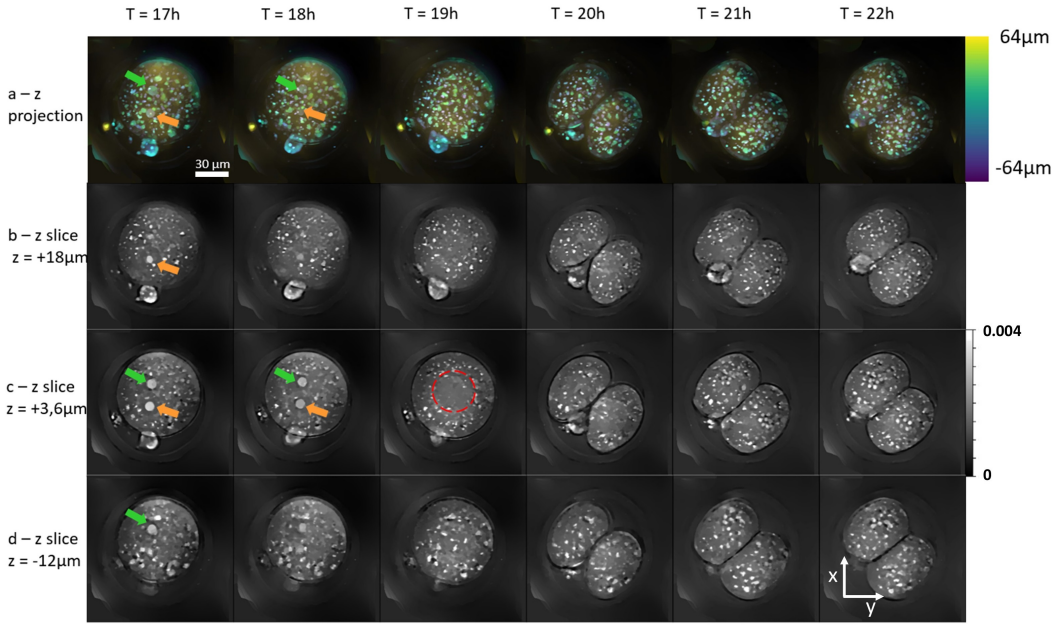


Figure 3.24: Time lapse of the first cell division of a mouse embryo. Scale bar is 30 µm.

On Fig.3.24, the pro-nuclei that are originally far away start to slowly converge toward each other before merging at t = 19 h. Thanks to the 3D reconstruction we can observe

that the pro-nuclei that were 20 μm away in the z direction at t = 17 h are visible on the same slice (Fig.3.24 c - $z = +3.6\ \mu m$) at t = 18 h. This is an indication that the pro-nuclei movement is truly in 3D, and that our prototype can well positioned and follow them in 3D. After the pro-nuclei have merged at t = 19 h (red dashed circle), the zygote divides in a two cell embryo at t = 20 h. The pro-nuclei fusion is an event that can be reliably used to predict the incoming first mitosis of the embryo.

Another example of the ability of our setup to locate accurately object in 3D is shown on the 4 cells stage, at t = 51 h on Fig.3.23. The polar body is at the bottom of the embryo and appears on the $z = -12\ \mu m$ slice (red arrow) but it is absent of the $+18\ \mu m$ slice. This highlights the interest in using 3D imaging when the sample to study has a 3D behavior.

Seventy hours after the beginning of the time-lapse (t = 81 h), a cavity (called the blastocoel) starts forming on the top of the embryo. It is important to notice that the appearance of the blastocoel can be accurately identified thanks to the 3D images. In contrast, in z projection, the appearance of the blastocoel is hidden and not detectable. Later, at 104 h, two cell types become apparent: internal spherical cells forming the internal cell mass (white arrow in Fig.3.23) and tightly joined flattened cells at the periphery of the blastocyst (the trophoblast cells).

This 3D video microscope is a precise instrument that produce high resolution images over various key part of the embryo at multiple stages of development. Those 3D time lapse reconstructions are a first step toward a fully autonomous and non-invasive microscope to monitor the growing and healthiness of embryos. Importantly, these first results in collaboration with the CHU-Grenoble laid the foundation for a further grant opportunity that got funded in 2021 French National Research Agency (ANR), project LIVE 3D_CNN (ANR-21-CE19-0020).

3.2.7 3D time lapse of liver organoid

After this extensive study on embryo developing we wanted to focus on an other living multi-cellular organism, namely mouse liver organoids. The goal was to validate the

capabilities of our system on a different biological sample. Other biological sample are presented in Appendix B, this section focuses on organoids. Their existence relies on recent breakthroughs in the development of culture conditions for adult stem cells. Those new culture conditions enable long-term expansion of adult primary tissues from liver into self-organizing 3D structures named 'organoids'.

For biologists, such self-organizing 3D structures can be used to study genetic diseases as well as cancer in a more efficient way than what can actually be done with standard 2D cell cultures. That is why organoids can provide more accurate understanding of cell-cell communication, naturally improving the efficiency of $invitro$ studies. Moreover, these 3D structures might be suitable to replace animal models in specific conditions, therefore decreasing the amount of animals needed for specific experiments. All the results here below were conducted in collaboration with Dr. Luca Valenti, from the Università degli Studi Milano, Fondazione IRCCS Ca' Granda Policlinico under the project REVEAL (European Union's Horizon 2020 research program under grant agreement N°101016726). The group in Milano was responsible for all of the organoid culture and development.

3.2.7.1 Organoids culture conditions

Liver organoids were obtained following a previously published protocol **??**. Human liver samples kept at 4°C in Basal Medium (Advanced/DMEM F-12 supplemented with 1% Penicillin/Streptomycin, 1% glutamax, and $10\ mM$ HEPES) before processing. Samples were manually minced and washed twice with $10 mL$ wash medium (DMEM high-glucose supplemented with 1% FBS, $100 units/ml$ penicillin and 0.1 mg/ml streptomycin). Sample was then dissociated by enzymatic digestion (basal medium with collagenase and dispase II $0.125\ mg/ml$, DNaseI $0.1\ mg/ml$) at 37°C for no more than 90 minutes to obtain an 80-100% single cell solution. Solution was then filtered through a $70 \mu M$ pores cell strainer and volume was increased to $50 mL$ with ice-cold Wash Medium before centrifuging at $300 g$ for 5 minutes at 8°C. Pellets were resuspended in $15 mL$ wash medium and washed twice with $15 mL$ wash medium and once with $10 mL$ basal medium (each time pelleting the material

by centrifuging at $300\ g$ for 5 min at 8 °C). Cells were then resuspended in Matrigel and seeded in $40\,\mu L$ Matrigel drops in $35\,mm$ µ-Dishes (Ibidi, DE). Drops were incubated for 20 minutes before being overlayed with $1\ ml$ isolation medium (basal media supplemented with 1X B27 supplement without Vitamin A, 1X N2 supplement, $1\ mM$ N-Acetylcysteine, $500\ \mu g/ml$ R-spondin1, $10\ mM$ nicotinamide, $10\ nM$ recombinant human [Leu]-Gastrin I, $50\ ng/ml$ recombinant human EGF, $100\ ng/ml$ recombinant human FGF10, $5\ \mu M$ A83-01, $10\,\mu M$ forskolin, $25\ ng/ml$ recombinant human HGF, $25\ ng/mL$ noggin, $50\ ng/mL$ Wnt3a, 10 Y-27632). After 3-4 days, isolation medium was replaced with $1\ ml$ expansion medium (lacking noggin, Wnt3a and Y-27632).

3.2.7.2 Vibrations and solutions

For organoid imaging, the microscope making the subject of this thesis was moved to the laboratories in Milan and installed in their cell culture incubator. As detailed before, organoids were seeded in a viscous medium (Matrigel) surrounded by expansion medium, an environment very sensitive to the ambient vibrations. Indeed, in their laboratories, imaging vibrations were the main cause of artifacts during acquisitions. Three causes of vibrations were identified over the experiments conducted in Milan:

centrifuges [?] are placed on the same table as the incubator They often produce a high level of vibration.

The solution is of course to place the incubator in which the microscope sits on another table, free of centrifuges. But in a space-limited biological laboratory it is not always possible. In that case, for "short" time-lapse (less than 12h) it is possible to acquire them during the night when no one is working.

the expansion medium out-spaces the Matrigel: This time the organoids are in 8 well slides, where each rectangular hole is $9.4\ mm \times 10.7\ mm$.

the obvious way to reduce this source of vibration would be to use a different container or a different imaging setup. However, our goal is to cope with the constraints

that the biologists have, in order to keep exactly the same protocol as with our imaging setup. Moreover, the volume of expansion medium is not a true source of vibration but more of an amplification of the surrounding vibrations.

other people opening and closing incubators nearby. Even after finding a place without centrifuges, this space was still shared with other biologist also working with incubator. The "fast" closing of door was generating high peaks of vibrations that lasted only a few seconds. The solution here was purely social, one just had to remind other people to close or open the door softly.

Vibrations are intuitive to visualise on a video, while it is much harder on a static image. To give an idea of the intensity of the vibrations, we compared in Fig.3.25 images at different time-frames with the mean image from a video. Vibrations can have different effects on the image, e.g. it can blur it as we can see on the image at $t = 0\,ms$. The blurring is not the most problematic artefact that can happen, because it only destructs the high frequency creating a slight loss in resolution. What is more problematic is the displacement introduced by the vibration, like the image at $t = 150\,ms$ on Fig.3.25 where the roganoid shifts to the right. If every image is randomly shifted in every direction, the result of the reconstruction would be an average of every object (i.e. a blurred object) even if every raw image is sharp with well-defined high frequencies.

Even with the vibration issues acknowledged and dealt with, other problems were present on the raw images that decreased the quality of the reconstructed volumes. The different artefacts are detailed below.

3.2.7.3 Corrupted images to to high organoids concentration

When working with single objects alone in the field of view, it is easier to ensure that the incident plane wave on the biological object is effectively a plane wave (i.e. the intensity level of the background is close to unit). Moreover, when other objects are present they may interfere with the scattering of the sample under imaging. For the intrinsic culture

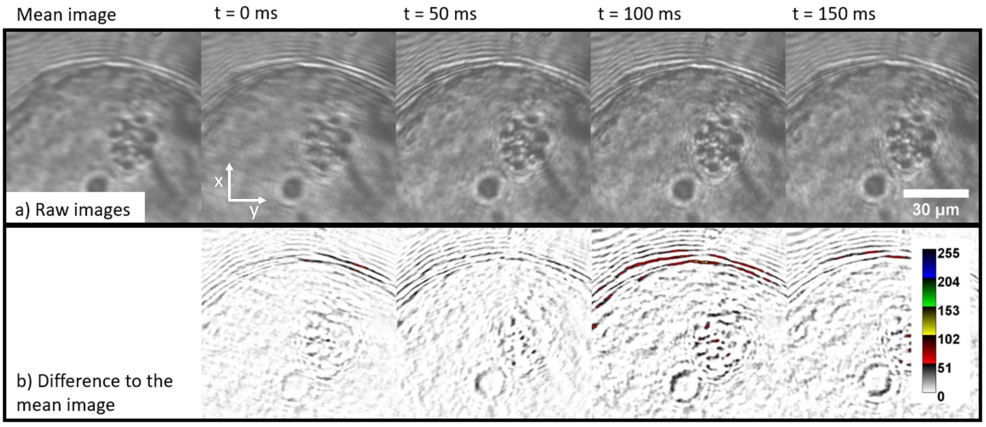


Figure 3.25: Visualisation of the vibration on organoids. a) is the raw images at various time during the video. The mean image is computed over the full 5 seconds of the video acquired. b) is the difference between the raw image and the mean image.

of organoids (cell seeding in Matrigel from pellet re-suspension), different samples can be easily find in the same field of view or overlay the one under imaging.

To have an idea of the effect of surrounding objects on the organoid under imaging, the normalised 84 images cropped around a liver organoid are presented in Fig.3.26. The on-axis images are clean without any perturbation, but many images (circled in red) present the "shadow" of an other organoid higher in the medium. This second organoid is detected during the angular change of the illumination. This shadow also skews the normalisation procedure, that fails to normalise two images with variable intensity levels of the background (6th row in Fig.3.26).

These corrupted images might interfere with the reconstruction algorithm. To qualitatively inspect the effect of such a shadow moving in the raw images, we performed the reconstruction twice. One with the full 84 measurements, one without the corrupted images, leaving only 54 measurements available for the reconstructions.

The major improvement that we can observe from removing the undesirable normalized images in Fig.3.27 is the increasing of the resolution. The small lipid droplet pointed out by the red arrow in Fig.3.27 d-h is completely blurred when the optimisation algorithm has to deal with the shadow passing over. Every interesting feature of the organoid is also

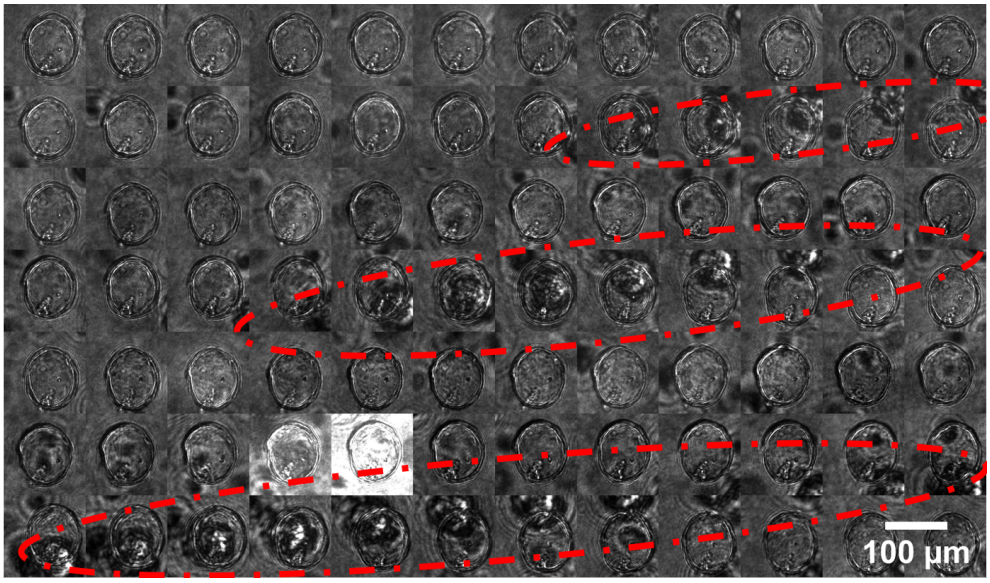


Figure 3.26: Cropped images of a liver organoid after normalisation for the first 84 LED's placed at a distance of 86 cm from the imaging holder. LED number increases from left to right, and from top to bottom.

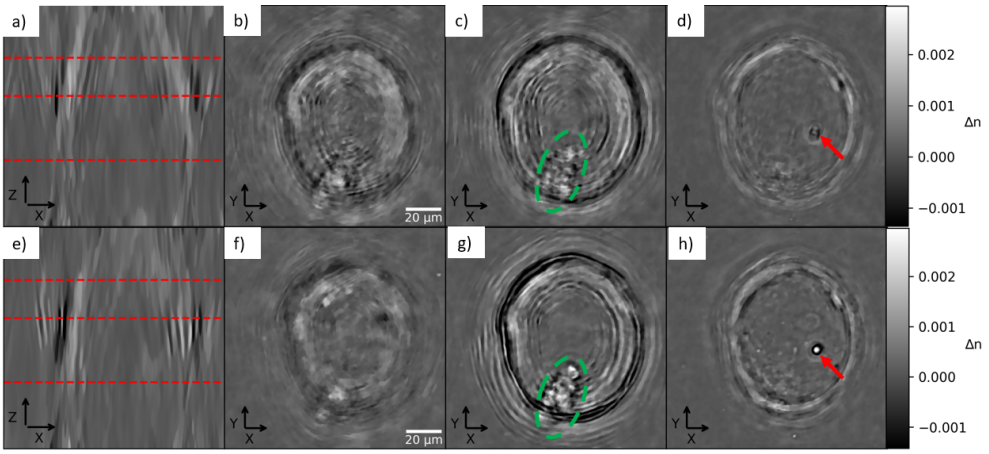


Figure 3.27: 3D reconstruction of a liver organoid with and without the corrupted images from Fig.3.26. a–d are slices for the volume reconstructed with all of the images (including the corrupted one - Sections are indicated with red dotted lines in a), from top to bottom. e–h are slices for the volume reconstructed with only the clean normalized images. b-f are slices at z = 29 µm. c-g are slices at z = 51 µm. d-h are slices at z = 87 µm. Scale bars are 20 µm wide

sharper, like the inner cellular mass circled in green (3.27 c-g). Finally on the first slice at $z = 29\ \mu m$ the amount of unwanted ring (circular structured) is lower and the back-

ground is smoother. Even though much improvements were made on the data selection and the vibration management, spherical aberrations due to the position of the organoid inside the Matrigel still corrupted the final reconstruction and defined outer border or inner structures of the organoids could not be retrieved yet (3.27 c-g).

This study indicates that images acquired with strong artefacts must not be used to perform the reconstruction. This is easy to deal with for a single reconstruction object, but for a time-lapse with hundreds of volumes this manual task can become time consuming and must be automated as in Section 3.2.7.6 with Fig.3.31.

The next Section investigates the effects of spherical aberration on the reconstructions.

3.2.7.4 Aberration due do the non-standard medium thickness

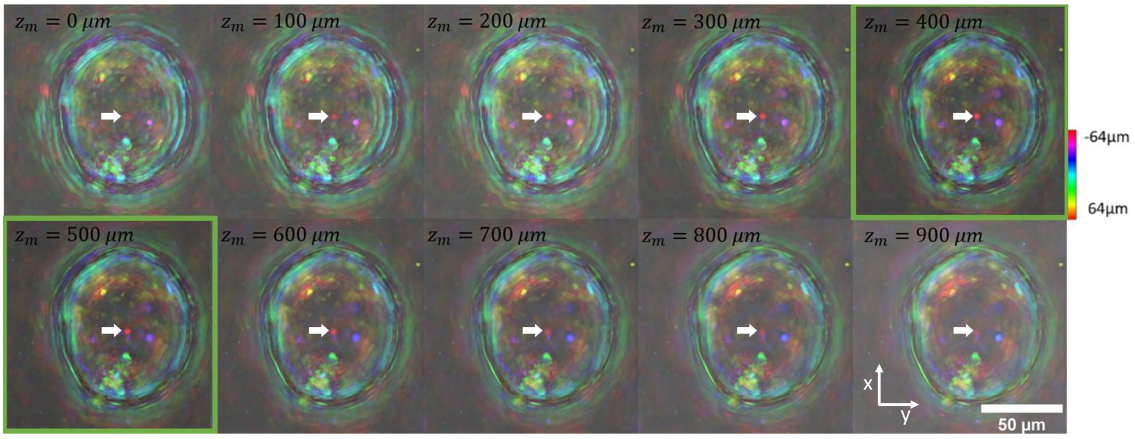


Figure 3.28: 3D reconstruction of a liver organoid projected in 2D with depth coloring for various aberration correction parameters.

The aberration study performed in Section 3.2.5.2 showed that a non-standard coverslip thickness or a non-zero medium thickness (the distance between the top of the coverslip and the object that is in focus) will introduce aberrations. In section 3.2.5.3 3.2.5.5, only the effect of the non-standard coverslip thickness was investigated since it was our major source of aberrations at that time. On the contrary, with the organoid setup the sample holder has a thickness that is corrected by the objective (i.e. $170\ \mu m$), but the samples are placed at various height, sometimes far from the bottom of the holder.

If this Z_m distance was known, we could directly use this value to have a more accurate forward model that would produce better reconstructions. In order to experimentally find this value from the measurements, we used a grid search on this parameter and performed a reconstruction with all the values ranging from $0\ \mu m$ to $1000\ \mu m$. Each reconstruction has been projected in 2D using a depth color code to produce Fig.3.28. The true shape of the organoid is not known, but in terms of refractive indexes we know that its outer layer should be constant and surrounded by a zero refractive index gap. This mean that the best object is the one with the less rings around the object and with the darkest background. With a focus on internal lipid droplet pointed by the white arrow in Fig.3.28, the dot is sharp only for values between 400 and $500\ \mu m$. The images with the green rectangle around seem to give slightly better results than without any correction at the end of the reconstruction process.

This study suggests that the organoid is floating between 400 and 500 µm above the bottom of the culture container, leaving a lot of space for other organoids to move below or on top of it. That might explain the issues that we were facing in section 3.2.7.3.

3.2.7.5 Time-Lapse of a single liver organoid at early stage of development

Imaging of organoids turn out to be challenging because of the presence of vibrations, high organoid concentration, and strong aberrations. Once all of theses issues were settled, the goal was to perform a 3D video of a growing liver oragnoid. The acquisition were performed with 84 illumination angles every 5 minutes for 6 hours. The reconstructions of theses 76 3D objects were performed with the Fast proximal gradient descent algorithm detailed in section 3.1.5.2.

Fig.3.29 shows the first steps of a liver organoid development. The white arrows point at small internal cavities that can be visualized over the time-lapse. Their number grows to 5, before they start merging to create a bigger one that is visible at 220 min on top of a small one. When looking at the purple color, which represents the lower part of the organoid, we can see that the cavity seems bigger in the lower part at the beginning (20

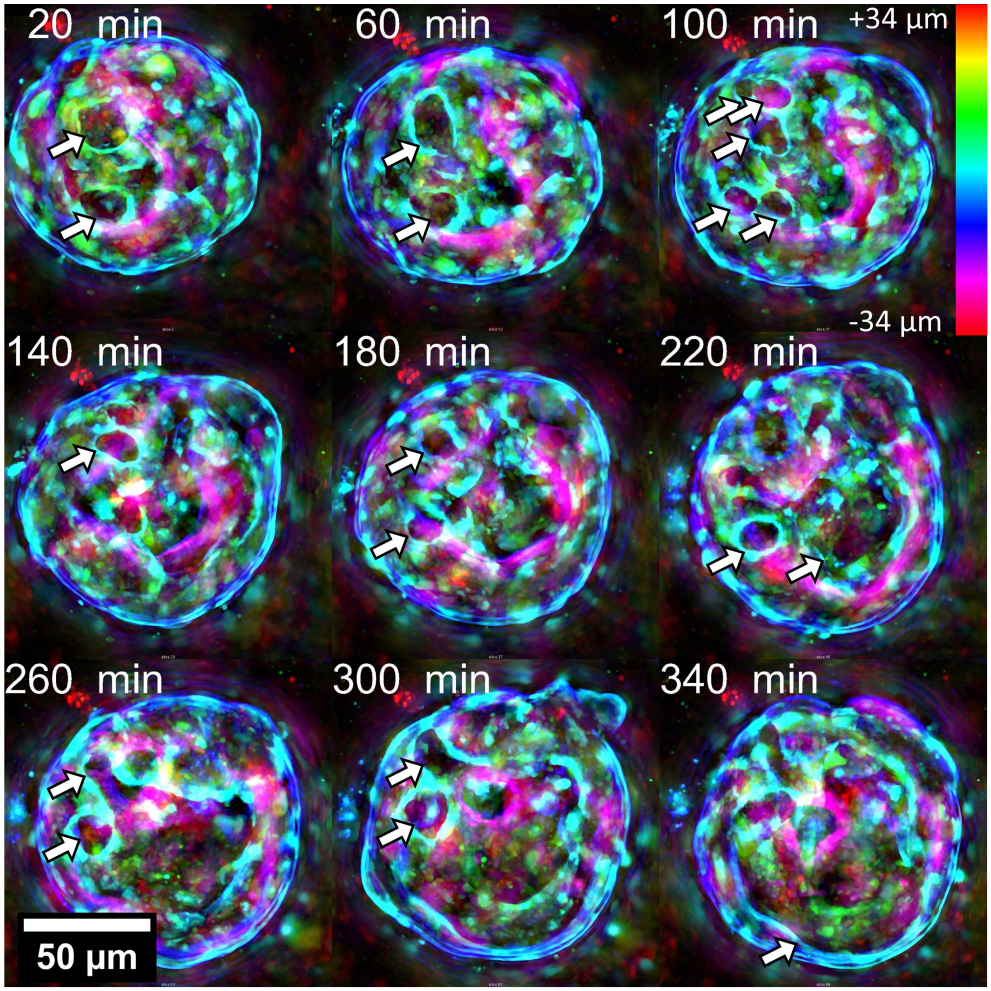


Figure 3.29: Time lapse visualization of the 3D reconstruction of a liver organoid projected in 2D with depth coloring.

and 60 min). Then it shrinks during a few hours to become small ($10\,\mu m$) in the lower part. This happens at the same time that the higher cavity grows, suggesting that this is just an internal movement of cells.

Only 9 volumes are shown on Fig.3.29, but the full video available in supplementary helps to understand the internal movement with a much higher frame rate (one every five minutes). This organoid was slightly moving in the xy direction for the first hour then stayed almost static.

3.2.7.6 *Time-Lapse of a moving liver organoid*

After further discussion with the biologist team, the organoid from section 3.2.7.5 did not appear to be very healthy as the cavity should be much bigger, creating an almost empty organoid that has the shape of football. Most of the cells are normally placed on the edge creating this eternal membrane that is characteristic of these liver organoid. This time we wanted to image a more healthy organoid that develops well.

A second time lapse was performed over night, to significantly reduce the vibration issue. The acquisition was performed with 84 illumination angles every 5 minutes for 12 h with the red LED array at $\lambda = 0.625\ \mu m$. The reconstructions of theses 146 3D objects were performed with the fast proximal gradient descent algorithm detailed in section 3.1.5.2.

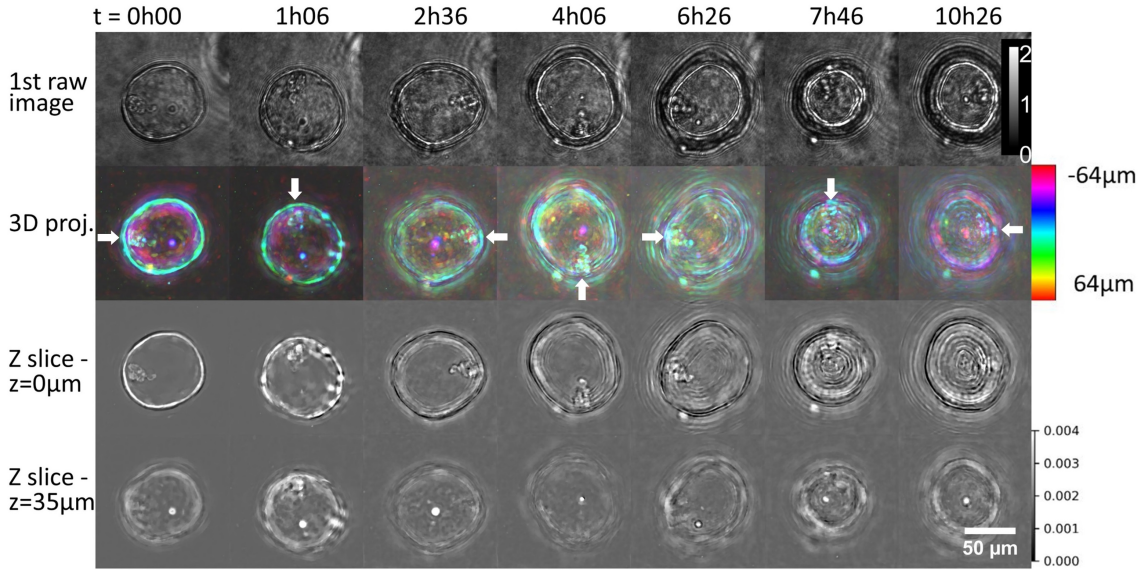


Figure 3.30: Time lapse visualization of the 3D reconstruction of a liver organoid. The first raw shows the acquisition image for the first LED only. The second raw shows a 3D depth projection of the reconstructed volume. The third raw shows a slice at z = 0 µm (in the center of the organoid). The last raw shows a slice at z = 35 µm (in the center of the organoid). The volume has been extracted every 90° turn.

Even though time-lapse are meant to be visualized as movie, seven interesting volume out of the 146 reconstructed were chosen to provide a static 2D representation of the key moment of this organoid development in Fig.3.30.

The first interesting behaviour that the biologists could observe for the first time thanks to the temporal nature of the reconstruction, is that the organoids are rotating around their center. On Fig.3.30 one volume has been extracted every 90° turn. The speed of rotation varies between 1.3°/min and 3°/min. All the small organoids (smaller than $100\ \mu m$) present in the field of view (results not shown) seemed to be rotating in the same direction, until they became larger and stop rotating. The rotation can be visualized from the inner cells that are attached to the membrane, clearly visible on the z slice at the center of the organoid, 3rd row on Fig.3.30. Another interesting behaviour is the temporary shrinkage of the organoid. As the time passes the organoid gets bigger and bigger, it can grows up to $300\ \mu m$ in a day. But sometimes the organoid stops growing and shrinks for a few hours before resuming its growth, as it can be observed between 6h 26 and 7h 46 on Fig.3.30.

The organoid is moving a lot in its culture medium; as long as it stays inside the field of view it can be tracked manually in the X and Y axis, but a movement in the Z axis creates a change of focus. At the beginning of the time-lapse the organoid is perfectly focus at its center, but as the time passes the defocus increases. This displacement along the Z axis increases the aberration due to non-optimal imaging condition. Using the technique from Section 3.2.7.4 we can estimate the medium thickness that induces aberration. This manual operation is time consuming and cannot be performed for every object from the time-lapse. This distance was computed for 5 time frame, and was estimated with a constant then log function for every other time frame of the time-lapse, shown in Fig.3.31

Once the depth of every object is known, this value can be used to take into account the aberration introduced by a non-standard medium thickness as introduced in section 3.2.5.2. Every volume was reconstructed again with this aberration correction procedure, and the comparison with and without correction are visible on Fig.3.32

The first improvement that our new model brings is the reduction of the diffraction pattern in the reconstruction. This is visible at 6h 26 on Fig.3.32, there are less parasite circular rings on the corrected reconstruction. Even if the corrected image is slightly better, it is still very blurry and a lot of diffraction pattern are present in the reconstruction. If

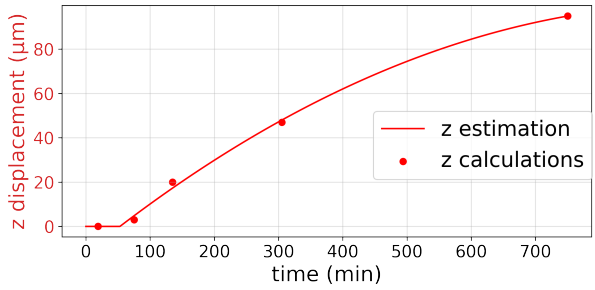


Figure 3.31: Z displacement of the organoid over time. The 5 dots represents the manual estimation of the z value. The solid line represents the function used to estimate the z displacement for every time an acquisition was performed.

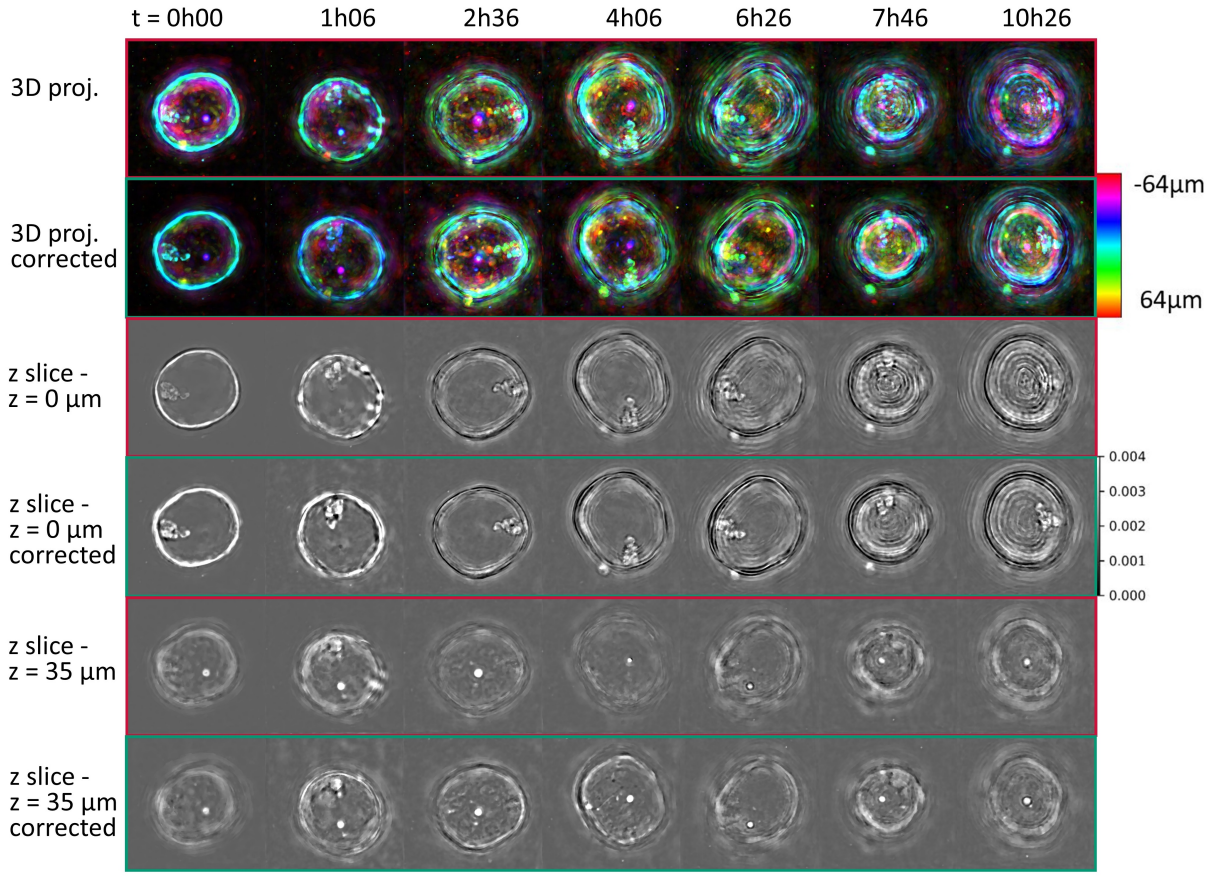


Figure 3.32: Comparison of 3D reconstruction of a liver organoid during a timelapse. Raw 1-2 are projection in 2D with depth coloring. Raw 3-4 show the central xy slice of the organoid. Raw 5-6 show a top xy slice of the organoid.

we compare the organoid of t=6h26 to the one from t=0 it is clear that we have worst results after a few hours and that the spherical aberration is not the main problem. The main difference between the nice looking organoid and the corrupted one is the change of focus due to organoid movement. Apparently out of focus organoid are less accurately reconstructed by our model.

Concerning the value of the refractive index, for the early stage of development (before 2 hours) the corrected model seems to yield brighter images, reflecting an higher refractive index reconstructed. Considering the fact that the refractive index is almost always underestimated due to the Z elongation, an higher refractive index is a good sign. The inner structure that allowed us to measure the rotation behaviour of theses organoids also seem more detailed and slightly brighter. This might be a side effect of the reduction of the parasite circular ring.

Even though the correction does not have a strong effect, is it important to remember that it comes with almost no extra computational cost, the only cost is an extra point wise multiplication that is neglectable compared to the hundred of FFT's that are performed at each model evaluation.

Most of the artefacts present in the reconstruction only appear when the imaged object is out of focus. On simulation, if an object is on focus or out of focus, the reconstructions will be equivalent. It appears that is not the case for images with our prototype. The explanation for this problem would be the presence of aberration inside our imaging system. The first step would be to characterize the pupils function of the system as in [90].

3.2.8 Limitations of the imaging setup

3.2.8.1 Acquisition time

To date, the total acquisition time of 2 to 3 minutes (for 84 different angles) limits the application of the system to biological samples that evolve in the range of minutes to hours, e.g embryon, cell culture and organoids. In order to address the imaging of living samples

at a faster rate, one could use fewer and brighter LEDs as in [76]. A ring of 8 powerful LEDs is placed on top of commercial inverted microscope, enabling them to image in 3D at a 10Hz rate.

3.2.8.2 Sensitivity to the focus

Theoretically, the focus placement inside the object has no impact. If we do the focus at the center of the object, then the reconstructed volume will have the reconstructed object at its center. If we focus on the top of the object, then the reconstructed volume will be in the lower part of the reconstructed object. In practice, when we are not focus exactly in the center of an object, it creates artefact. The circular interference created by the defocus are not reconstructed properly and end up directly in the reconstructed volume (visible at the end of the time-lapse from Fig 3.32).

For the imaging of single object, this will be solved by using a xy motorized stage to follow the individual object and then use a piezoelectric stage to perform an auto-focus on the desired object.

To image multiple 3D object placed at different heights (and thus being out of focus), this sensitivity to the focus will have to be studied and corrected.

3.2.8.3 Lack of quantitative information

With a tight control on the experimental conditions the previous limitations are not problematic. But we chose by conception to have huge lack of information to perform the 3D reconstructions. We are missing the phase of the acquired light field because we are not using an interferometric system for simplicity. We are also loosing a lot of information because we have a limited angular illumination due to the low NA objective we use to maintain a wide field of view. All of this missing information results in three major issues.

First we are missing the low frequency of objects. Most of the signal we acquired with our CMOS sensor comes from the edges in the observed object, leading to reconstructed object that are not correctly measured in terms of refractive index (i.e. objects with con-

stant n are not filled, but only the edges are visible). To fill the object we rely mainly on regularisation with the TV penalisation.

Then the low angle coverage results in an axial elongation of object as discussed in section 3.1.4.3. With recent method like deep-learning, this issue can be addressed with learning based methods as presented in Chapter 4.

Finally this elongation reduces the reconstructed refractive index, leading to a non-quantitative system concerning the value of the refractive index for every voxel.

Appendix B

Appendix on 3D biological reconstruction

The microscope presented in Section 3.2 was used to image many other biological sample than embryo and organoids. For wider object the field of view of $200 \times 200\mu m$ is not enough. To image with a wider field of view, we decided to use a lower magnification objective (10X instead of 20X), with a lower NA (0.2 instead of 0.4). This should decrease the resolution by 2.

To increase the field of view once more without degrading the resolution, we decided to use a shorter lens tube of 9 cm instead of the nominal 18cm. This reduces the magnification from 10X to 5X while maintaining a 10X NA and resolution.

B.1 IMAGING WITH WIDER FIELD OF VIEW

The first sample to be used with a wider field of view is challenging by its support. The imaged section in Fig. B.1 is a part of a microfluidic channel in which organoid vascularization is studied.

The second sample in Fig. B.2 is a pancreas organoid studied for its possible insulin generation in presence of glucose.

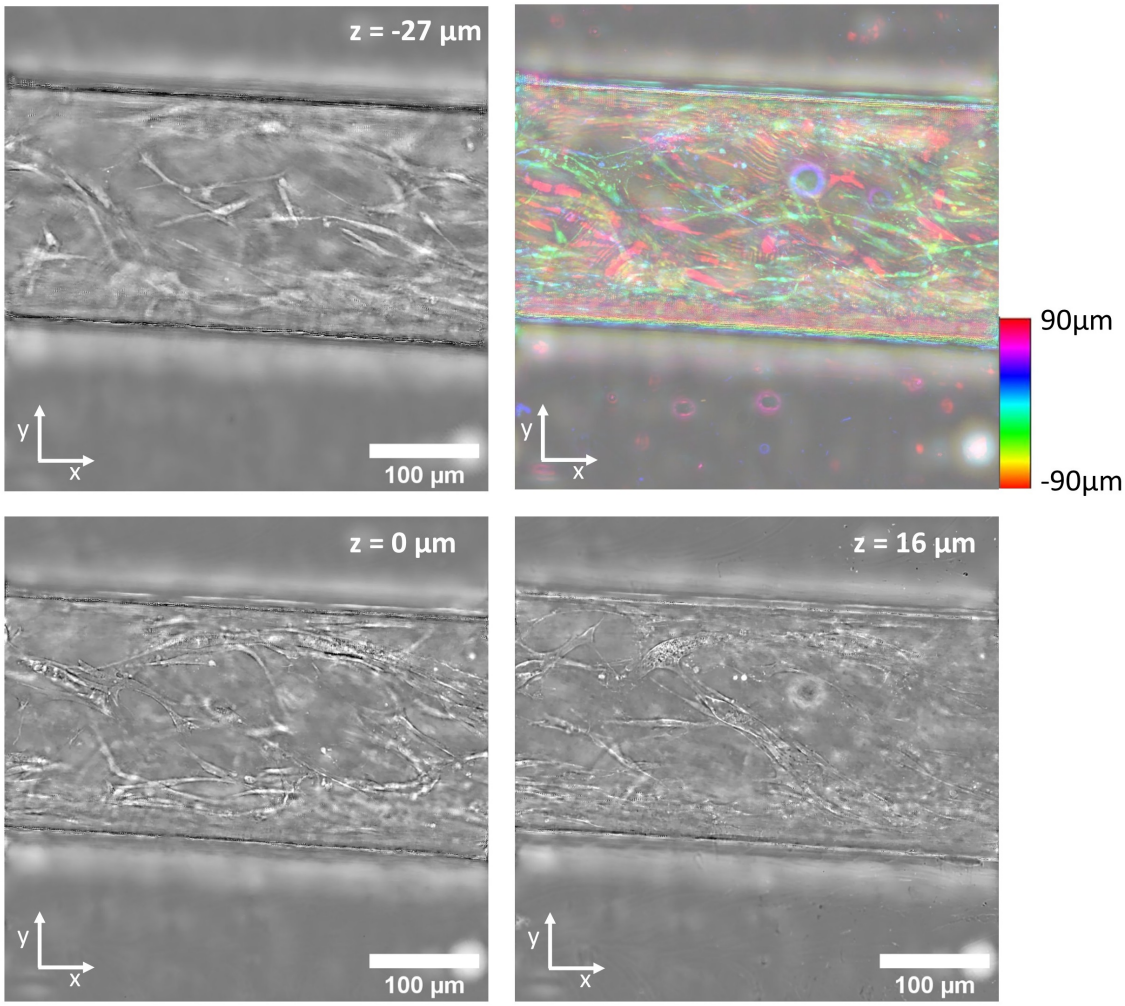


Figure B.1: xy cross section of a microfluidic channel growing a vascular network

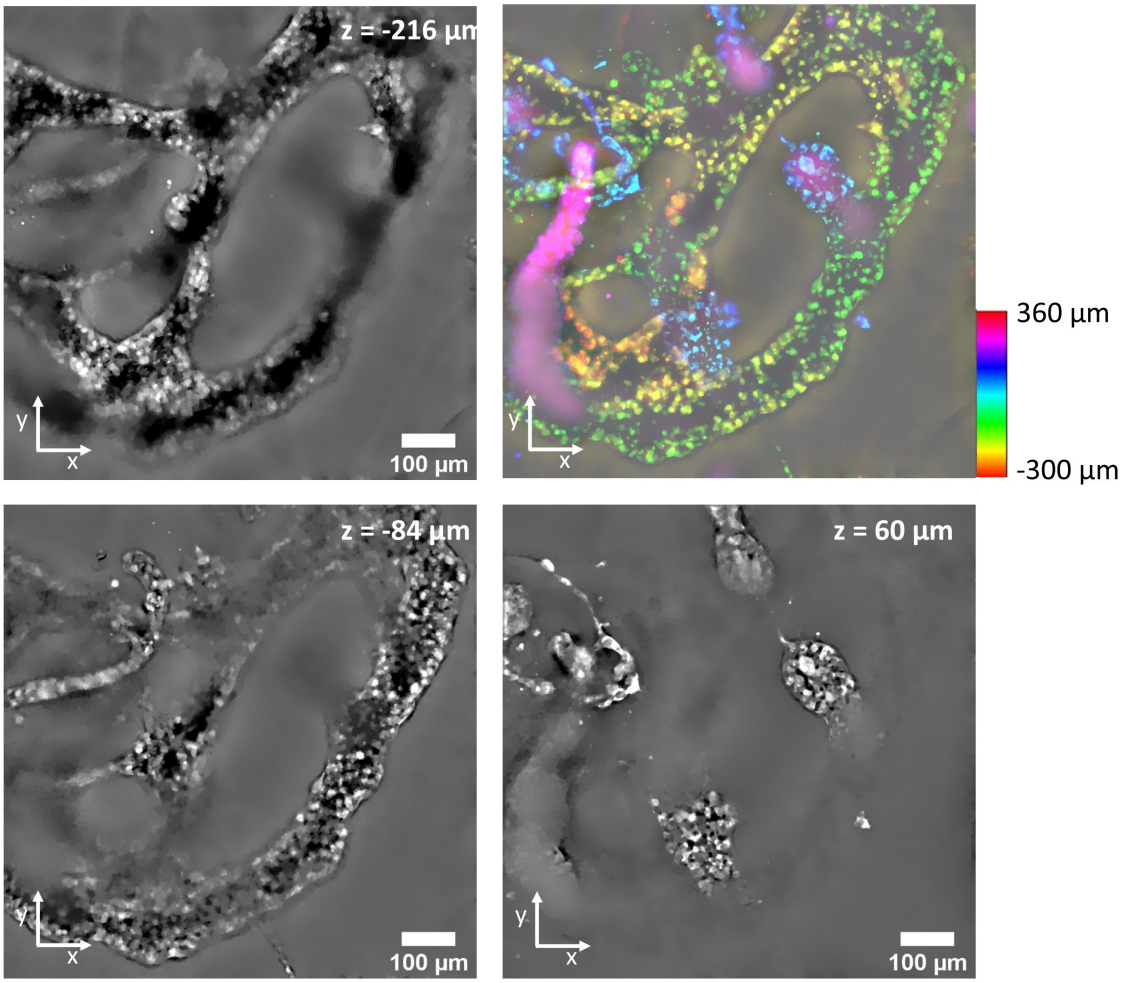


Figure B.2: xy cross section of a pancreas oragnoid

Chapter 4

Deep learning and Intensity Diffraction Tomography

The goal of this chapter is to introduce deep learning, starting with its short history before detailing the main parts that make it so popular in computer vision and in biological imaging. Then it will be used to solve our main 3D reconstruction artefacts. Importantly, results from this chapter were filed into a patent that is still under processing and orally presented on [98]

4.1 DEEP LEARNING: BASICS

In this section, the deep learning essentials will be introduced, with a focus on convolutional neural networks, which are the foundation of modern computer vision algorithms and computational microscopy.

Artificial intelligence (AI) has experienced several hype cycles, with alternating periods of reduced funding or high interest in AI research. Thanks to the increasing in power compute and data available, a part of AI, deep learning, exploded during the last decade. Year after year it is now making huge breakthroughs in many fields. It started with computer vision, but quickly extended to machine translation or speech recognition, as examples

[72].

To have a better idea of how fast deep learning (DL) improved the performance in image recognition, let's have a look at the ImageNet [24] image classification competition that started in 2009. As shown on Fig.4.1 the level stagnated for the first few year with high error rate ($> 25\%$). From 2012 every winner was using deep learning architecture and the error rate quickly dropped bellow the 5% that is an estimation of the human error rate.

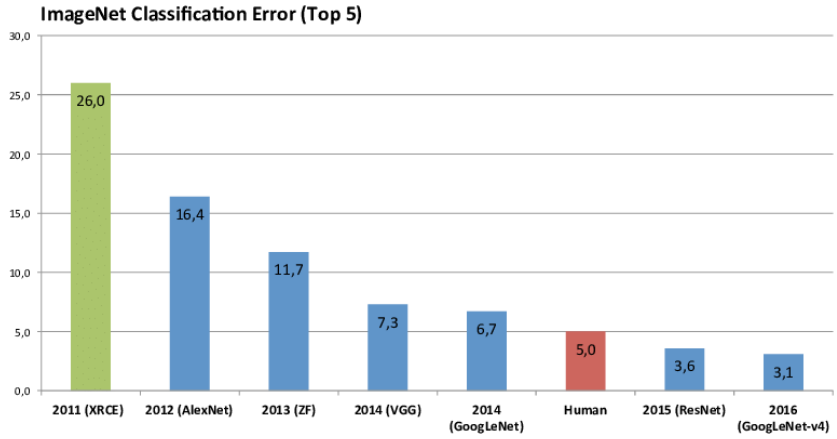


Figure 4.1: Performance of the winner of ImageNet [24] image recognition challenge over the years (von Zitzewitz 2017). Deep learning architectures are in blue. Fig. 10 from [82]

This first successful application quickly brought the interest from other fields where super-human performances were rapidly obtained, in visual reasoning [114], in poker [87] and traffic sign recognition [18].

4.1.1 Machine learning

An important part of AI is machine learning, which contains all of the algorithms that let a machine autonomously train its model to performed a desired task . At present, machines refers to computer programs and the implemented algorithms aim at estimating the probability densities that describes the data under observation. Every machine learning model follows the same principles.

A model $f(x, W)$ has a set of parameters W which influence its comportment. The model has to map all the input x to all the corresponding output y in a way that $y =$

$f(x, W)$. We can consider that x is drawn from an unknown probability distribution $P(X)$. To build the best model possible, the conditional probability distribution $P(Y|X)$ is required but never available for real world examples. For real application only some inputs χ and their matching output Υ are known. Machine learning goals is to find the best $\hat{W}$ with the intent that the predicted $\hat{y} = f(x, \hat{W})$ will match as closely as possible the real y. If everything works perfectly, the known input data χ is well modeled as well as the true underling distribution. We can then say that the model did not remember all the examples, but has truly learned the unknown data distribution. Another way to look at the problem is to see it as an optimisation problem on a supplied dataset (χ, Υ). Even-though it is not always possible to map a machine learning to an optimisation problem, it will be the case for all the model presented in this document.

The two major goal of machine learning are the following:

- 1 - Searching for the best model. Since there are countless way to create a model, a great amount of research focus in improving the existing model or building new ones.

- 2 - Estimating the best W possible.

Concerning the weight estimation, the gradient descent is by far the most efficient and democratized method. First the idea is to calculate $\hat{y} = f(x, W)$ on the training examples and an initial W and then estimate the error between $\hat{y}$ and y with a loss function $L(y, \hat{y})$. The loss function measures the similarity of $\hat{y}$ and y and if the loss was correctly build, smaller values of $L(y, \hat{y})$ shows higher similarity. Then by computing the first and rarely the second derivatives of $L(y, \hat{y})$ with respect to W, a small adjustment dW is computed such that $L(y, f(x, \hat{W} + dW)) < L(y, f(x, \hat{W}))$. This is the gradient descent of a loss function and even if many alternatives were tested, it is currently the most efficient method for training (i.e. finding the best parameter W) for neural networks and multiple other machine learning models.

4.1.2 Neural networks

Deep learning is the part of machine learning where the model used is an artificial neural network. This artificial neuron was inspired by the biological neuron, and simply represents a transfer function where the inputs data are multiplied by coefficients and summed. This sum is then transformed by an activation function to produce the output. The coefficients are the weights, or trainable parameters of this model, and the goal is to find the weights W that produce the desired output for a given task and data.

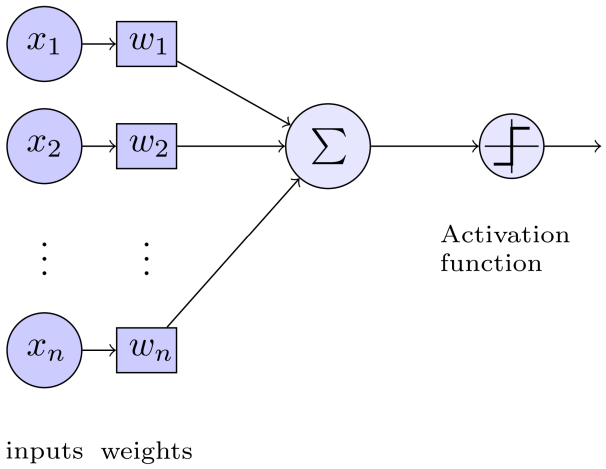


Figure 4.2: Illustration of an artificial neuron: the perceptron

This artificial neuron was named perceptron [105] and works as a linear regression, therefore it cannot solve complex non-linear problems. The multi-layer perceptron [108] combined with non-linear activation function is complex enough to guarantee that weights W exists for every imaginable task [50].

To find the desired weights for a given task, multiple steps are required and summarized in Fig.4.3. First, a forward pass is performed where each neuron receives the input data and computes the output results y_{pred} steps after steps. During the training the desired output y_{true} is know, and an error between the produced output and the real output can be computed through the loss function.

This error is then back-propagated layer after layer to compute the gradient of the

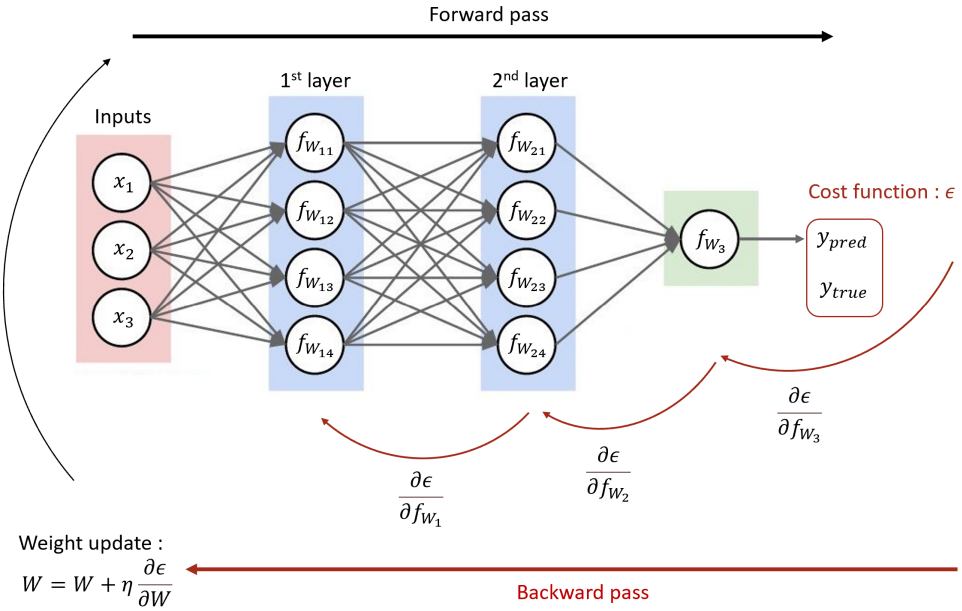


Figure 4.3: Schematic view of an artificial neural network composed of 2 hidden layers and its training loop composed of 4 parts. The forward pass computes the output. $epsilon$ is the cost function to compute the error of this output. The backpropagation pass computes how the weights should change to minimise the error. The actual change of the weights decreases through gradient descent.

weights relatively to the error. From this gradient, the weights can be updated. For more detail about the gradient descent please refer to Chapter 2 on optimisation (section 3.1.4).

4.1.3 Dataset

The gradient descent is performed on the training dataset, and the weights of the neural network will be fitted to have the lowest cost function possible. To control over-fitting issues (detailed in section 4.1.7), a validation dataset is commonly used to control the performance of the neural network in an unbiased way. Each model has intrinsic hyper-parameters that can be tuned, such as the number of layers, the learning rate of the optimizer, the type of the optimizer etc. This tuning is normally realized on the dataset itself, therefore the performance of the model on this dataset will not reflect the overall true performance. A third dataset is normally used to estimate the true performance of the model on unseen data, the test dataset. In theory we assume that these three dataset are sampled

from the same data distribution, but in practice we often have only one dataset that we split into training and validation. The validation is used as the test dataset in many cases, even if the predicted performance will be biased.

4.1.4 Convolutional neural network

The first work on convolutional neural network (CNN) started in 1980 [38] and stay quite confidential for thirty years until this field of research started to explode. The idea behind CNN is brilliant: instead of applying fully connected layers to a (long) 1D array of 2D points that automatically hides the structural information of the dataset (i.e. the interconnection between points), convolution operations are used.

4.1.4.1 Convolution

One may ask why it is an issue to apply fully connected layer directly between the input and the output of our data.

The first major drawback of fully connecting layers for image treatment is the absence of translation invariance, for example in the case of an image translation task where empty circles need to be filled. Once a network was trained to fill circles on the top left of an image, it will not be able to fill circles anywhere else on the image. Thus the training data need to contain examples of circles everywhere on the image. For a simple task such as circle filling this is tedious, but doable. However, for a more complex biological sample, it might not be possible to gather images with all the possible positions and orientations of an object or detail.

The second major drawback is the memory consumption of fully connecting layer for an image-to-image task. For a 512×512 image, there are 262144 pixels. If x and $\hat{y}$ have that 1D shape, the weight matrix W of a single perceptron would have more than 68 719 millions parameters. With float 32 bits, it represents more that 256 Gb of memory, which are at the current time an order of magnitude bigger to be used on any commercial GPU.

The solution to both issues is the use of discrete convolution with a kernel. The parameter of these convolution kernel are the trainable parameter. The output of the convolution is a feature map, that is then feed to the activation function. The number of parameter is given by the size of the kernel and does not depends on the size of the image. This provides a good scaling ability without hardware bottleneck, as well as a native 2D or 3D data support without any 1D vectorization. The translation invariance comes from the definition of a convolution that apply the same operation everywhere on an image. That means that once the network can fill a circle at one place of an image, it can perform this change at any position.

4.1.4.2 Receptive field and pooling

The use of convolution is not perfect either, even if it is widely use. The main drawback comes from the fact that an output pixel only has information from a part of the input image, depending on the kernel size and the presence of pooling layer. The number of pixels available from an output is called the "*receptive field*". A good network architecture must have a "*receptive field*" that is big enough to correctly perform the desired task. For the example of circle filling, it should be big enough to "see" the whole circle before filling it. An easy way to increase the receptive field is to increase the kernel size, but the number of trainable parameter increases quadratic with the kernel size. That is why most of the architecture uses 3×3 or 5×5 convolution kernel.

Another way to increase the receptive field is to put multiple convolutions one after another, thus allowing to linearly increase the receptive field. By stacking two 3×3 layers, the output can have information from a 4×4 image on input. This is why a lot of convolutions are required to have a big receptive fields. This number can be reduce by using pooling operation. The pooling operation aims at selecting only the most important information, thus reducing the amount of computing power required. The most used pooling is the max pooling that selects only the highest value, disregarding all the other values. Once the high level features have been extracted by the accumulation of convolution, activation,

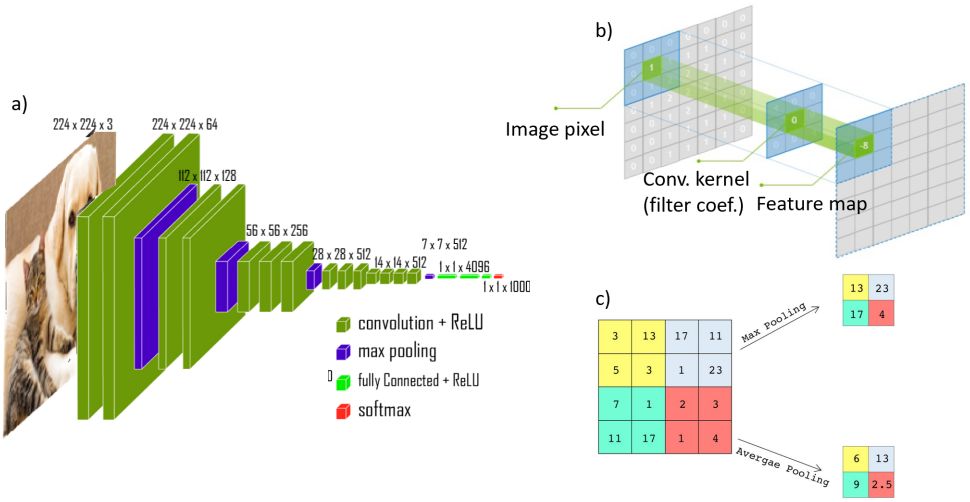


Figure 4.4: a) Architecture of a deep neural network. b) Convolution mechanism. c) Max and average pooling example. Images modified from https://blogs.nvidia.com/blog/2018/09/05/

pooling, the following part can be shaped to match the desired output structure.

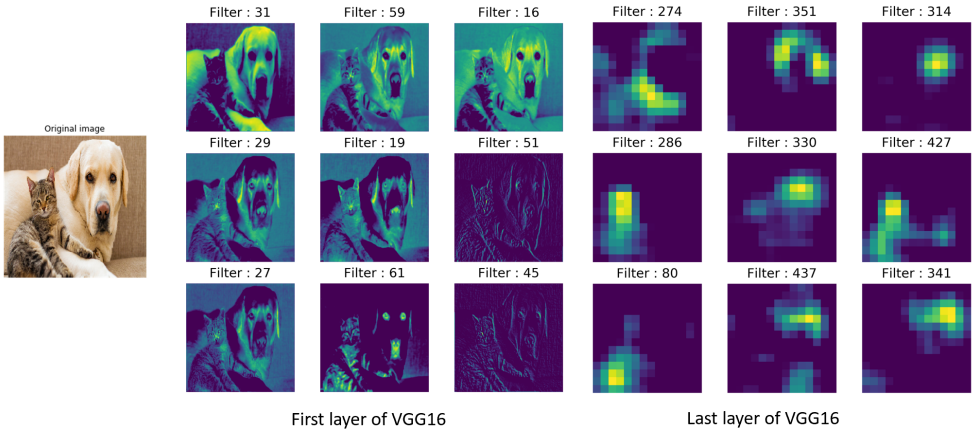


Figure 4.5: Feature map of the VGG16 with a cat and dog input. Genarated with code from https://gist.github.com/TejaSreenivas

In the case of VGG16 presented as an example in Fig.4.4, it is used to classify images into 1000 classes, therefore the output is a 1×1000 1D vector. After each convolution block, the image can still be visualized, for example with a dog and cat in Fig.4.5. In the first layers the filters are reasonably closed to usual filters, such as for high or low frequency selection.

Deeper inside the network, the filters are becoming more complex and less intelligible. Nonetheless these advanced filters are decisive in the process of correctly extracting the features needful to perform correctly the assigned task.

4.1.5 Image to image network and U-net

The classification network presented in section 4.1.4 outputs a 1D probability vector, but if we want an image as an output other architectures need to be use.

For these image-to-image task, the U-net [104] architecture has been widely used, after being introduced in 2015 for biological image segmentation. It combines many interesting architecture choices that made it a reference in the domain. First there is a compression stage where the input image is down-sampled with convolution and pooling to extract the meaningful feature present inside the image. As shown on Fig.4.6 the features extracted have multiple resolutions,and the pooling operation reduces the resolution by a factor of 2 at each convolution block. This gives room to increasing the number of features at each block with an increasing number of convolution channels. Moreover, from each condensed image representation there is the up-sampling branch that goes back to the original image size. For up-sampling, transposed 2D convolution are used, where a kernel of trainable parameters perform the size expansion. This auto-encoder architecture helps correcting the image in its low frequencies, but for higher frequency details skip connection were introduced. The data are just copied and concatenated with the result of an up-convoluted layer that has the same shape. The goal is to keep the features that were extracted by the compression filter to help the up-sampling part in keeping high frequency details. The U-net was originally used for classification on every pixels (segmentation), this is the reason why it uses 1×1 convolution to avoid the downside of fully connected layers. For image-to-image translation, slightly different final layers are used. The sigmoid that is almost always used for classification was removed and replace by a ReLU [37].

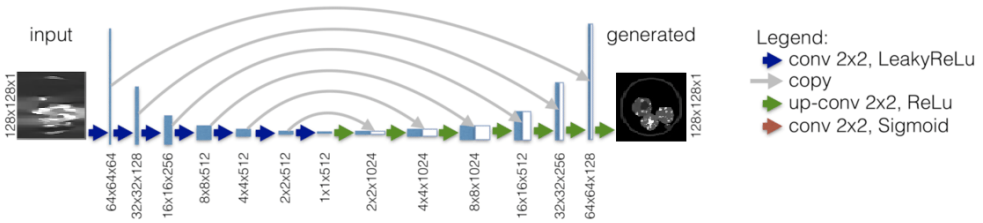


Figure 4.6: Modified U-net architecture.

4.1.6 GPU and training

Of course the new model architectures and good practices around deep learning helped a lot to increase its performances, even though the main improvements came from the hardware. With the increase of compute power thanks to the modern GPUs (developed first for gamers), the number of operations that can be performed is ten to hundred time higher than a recent CPU that is already ten times faster than a CPU from the early 2000. GPUs are faster for deep learning training because they have thousands of cores that can parallelize the operations, while a CPU only has between 4 and 32 cores. The CPU is more robust and can crunch a lot of various task, but to perform a task that can be highly parallelized, like huge matrix product or a high number of convolution, GPUs are faster. In addition, the price of data storage also dropped rapidly during the past two decades allowing researchers and private companies to gather, label and store huge amount of data (see 4.7).

4.1.7 Overfitting

One of the many way a neural network can malfunction is overfitting. This issue is not specific to deep learning but can be found in many machine learning task. It can be naively illustrated with polynomial fitting on a 1D regression task in Fig.4.8.

If the number of parameters is too small and the model is not complex enough to represent the underlying data distribution, we are in an underfitting situation. On the other side of the spectrum, if the model is too flexible it will match every data points of the train-

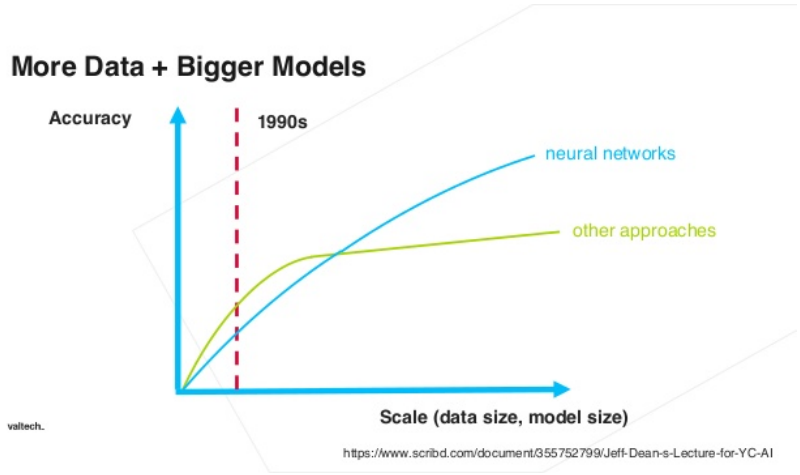


Figure 4.7: A simplified way to look at data-scaling of deep learning algorithms.

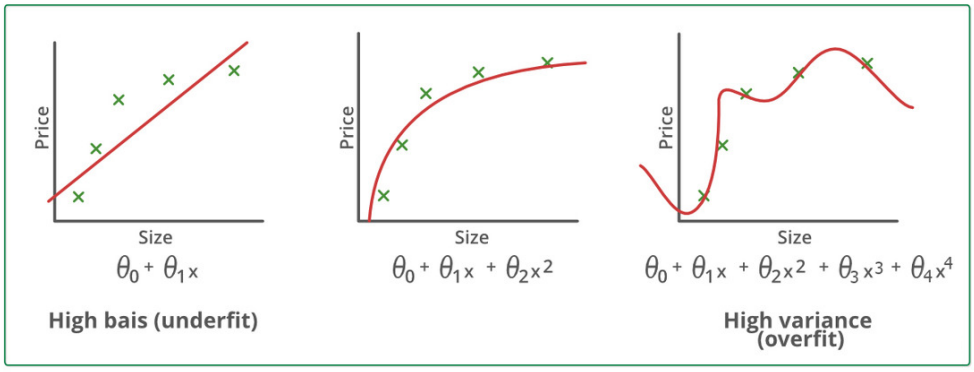


Figure 4.8: Overfitting visual explanation. Source: https://fr.acervolima.com/

ing set without understanding the underlying distribution, namely overfitting. The model does not learn the proper data distribution but remembers all the training points. Then, when it will face new unseen data, its predictions will be wrong.

Overfitting appears when the number of training data is lower than the trainable parameter of a model. Indeed, manually annotating data points is a tedious work and modern deep learning (DL) architecture often have ten's or hundred's of millions parameters. That is why overfitting is a common problem in the DL field.

In practice to detect overfitting while training, we need to use the validation dataset. By monitoring the error (loss) on the training dataset and the validation after each training step, one can notice if and when the model if overfitting. On Fig.4.9 it happens when

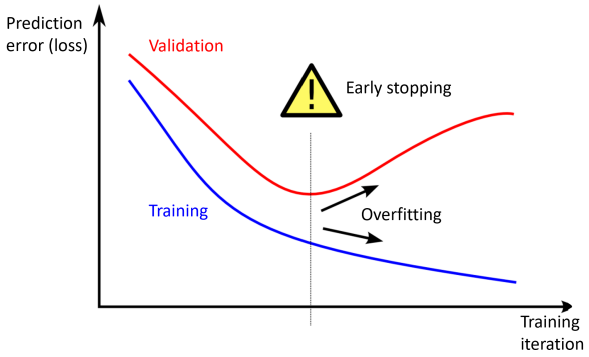


Figure 4.9: Overfitting detection and early stopping placement.

the validation loss starts to rise while the training loss keeps decreasing. A commonly used technique to deal with overfitting is *early-stopping* [11], where overfitting is tracked and the training stopped as soon as the model starts to overfit. The overfitting control can also be performed by increasing the training dataset, either by gathering more data (often expensive and tedious work) or by virtual data augmentation. On images, data augmentation will slightly alter the available data by modifying the original images with simple operations, such as cropping, rotating, changing exposure, adding noise etc. The idea is to alter the input data to prevent the network from remembering each image, and instead forcing it to find the important characteristics of images. With the same idea but for the inner weights, the dropout [123] is extensively used. With dropout random neuron are deactivated at each training stage, thus limiting the possible complex co-adaptations and forcing the network to find robust features. To accelerate and stabilise the training, the *batch-normalisation* normalises the data to the mean and standard deviation of the current batch [53] . The same image will have therefore slightly different pixel values depending on the overall images present in the current batch, also preventing overfitting.

4.2 DEEP LEARNING AND TOMOGRAPHIC RECONSTRUCTION FROM SCATTERING IMAGES

4.2.1 State of the art

The field of optical diffraction tomography was not unaffected by the deep learning wave and during the past five years various strategies have been used. A graphical view is presented in Fig.4.10.

Deep Learning (DL) has been originally used to pre-process raw images from the acquisition setup and remove noise [14]. This problematic is omnipresent with quantitative phase measurements, where the phase is obtained from the interference with a reference beam. Since we are working with intensity only images (see the previous Chapters), our images are much cleaner and such network is not needed. An other use of DL on measurements was to help removing the unusable measurements [110], whether because they are too noisy or because another object has perturbed the acquisition. This step has been manually performed in our acquisitions (See section 3.2.7.3).

A different research team trained a network to go directly from the normalized measurement to the final 3D map of the refractive indexes [60] [61]. This task is the most difficult of all, since it does not includes the physic of light scattering and therefore was mainly performed on phantoms and spherical objects. The variety in wavelengths, refractive index range, and experimental conditions, makes the amount and the variety of data required for real bio-usage gigantic. Even if this technique would be the fastest, at the current time no direct inversion with DL was performed on real, diverse, biological samples.

A more precise approach was to only introduce deep learning for the regularisation part with deep priors [140] [144]. The idea is to replace the common $L1$ or $Total\ variation$ regularisation by a network trained to perform a similar task. The main drawback of this technique was that at each iteration the gradient descent from the measurements was elongating and increasing the edges instead of providing compact and full objects. It can be

considered as an improved regularisation based on deep learning.

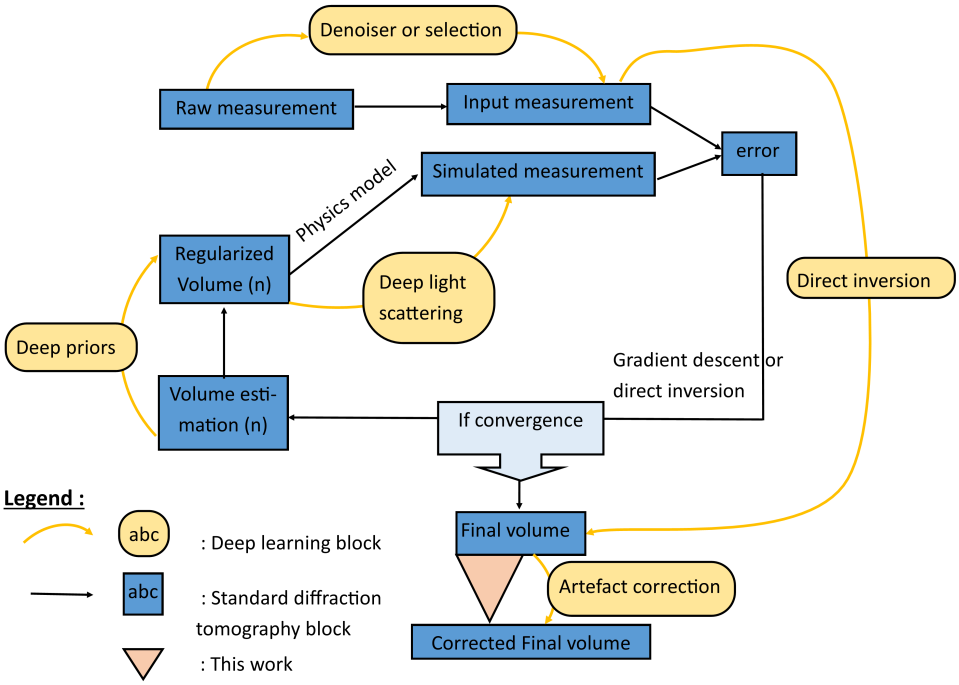


Figure 4.10: Integration of the deep learning strategies inside a standard learning tomography from transmitted images.

Deep learning was also introduced inside the light scattering model with a *"Physics-informed neural networks"* [112]. A modified U-net is trained to compute the scattered field induced by the 3D refractive index map. This trained network can approach the accuracy of a finite-element solving of the Maxwell equation, therefore being more accurate than multi-slice models even for high illumination angles. Once the network is trained, it takes only a few tens of seconds instead of tens of minutes to compute the forward model. This is groundbreaking for high resolution systems with high numerical aperture (NA > 1), where most of the error comes from the accuracy of the scattering model. However, in our case, this technique might not be relevant, since the multi-slice model is good enough to cop with low NA objectives and aberrating optical systems.

Finally, deep learning was trained to correct optical aberrations on volumes reconstructed with a linear model and direct inversion [79] [83]. For this approach, random

volumes are generated and reconstructed, and the network is trained on simulated data only to go from the corrupted volume toward the real ones. Lim *et. al.* [79] were the first one to implement such a method and restraint their use case to red blood cells only. This was the first step to show that training on simulation was good enough to generalize to real biological red blood cells. Then Matlock and Tian [83] trained their network on random stack of 2D images as 3D volume, and the network perfectly generalized to real biological samples without any fine-tuning.

Our approach detailed below also lies in the "artefact correction" category.

4.2.2 Artefact correction from random stacked images as 3D biological volumes

The approach that looked the most promising for our 3D reconstructions was the work done by Matlock and Tian [83], were at the starting point of the network was given the best possible linear reconstruction and it had to guess the real object from that. In their work, the axial resolution was truly improved and, since the network was train on random non-biological images, its generalisation capabilities are not questionable.

We decided to use the exact same method for the data generation and the training of the neural network, except that we replace their reconstruction model by our model. The goal for our data processing would be to decrease the elongation on the z axis and improve the resolution of n. Matlock and Tian have a high resolution and low aberration microscope (*i.e.* an expensive and big inverted commercial microscope) coupled with a linear model reconstruction, and the goal of their neural network is to correct their model error. On our side we have a multi-slice reconstruction model, but coupled with cheaper and more prone to aberration hardware. The sections below will describe the adaptation of their algorithms to our model.

4.2.2.1 Data generation from random images

In order to train a network to correct artefacts from reconstructed volumes, the first task is to create a training dataset with paired volumes with and without artefacts. In 2D phase

imaging, this kind of datasets can be acquired by taking an image with a reference microscope after imaging the sample with the current method [103]. But in 3D the few quantitative microscopes that exists are still laboratory prototype, that were not available for us to use. Note that their high numerical aperture force them to have a huge magnification, leading to a field of view too small for our applications.

Since we cannot get real training data for our artefact correction model, we will have to simulate the data. Simulated volumes of size $128 \times 128 \times 30$ were be generated by stacking 30 normalized images randomly drown and cropped from the COCO dataset [81] that contains over 300 millions images taking from everyday images.

The normalisation procedure follow the next steps:

- randomly rotate the image between 0 and 360°;
- divide the pixel values by a random number between 255*20 and 255*1000. This ensures a maximum refractive index gap between 0.05 and 0.001.
- 30% of the time return an image with 0 for every pixel. This is to train the network to return empty refractive indexes where there is no object;

From this generated volume, the microscope intensity images are simulated to approach what we would have measured if the generated object was place inside our microscope. From those simulated measurement, we reconstruct the object using the same exact loop that the one used in Chapter 2 for the time lapse reconstruction.

An example of a generated and simulated volume is shown in Fig.4.11. The axial (*i. e.* Z axis) elongation in clearly visible in the YZ slice. This elongation is reflected on the XY slice by the fact that the shape present in one slice only appears on the previous and next slices, like the brick wall or the surfer.

2800 volumes have been generated and simulated, providing more than 80 000 paired images for training the neural network. 200 more volumes have been generated to create a validation dataset.

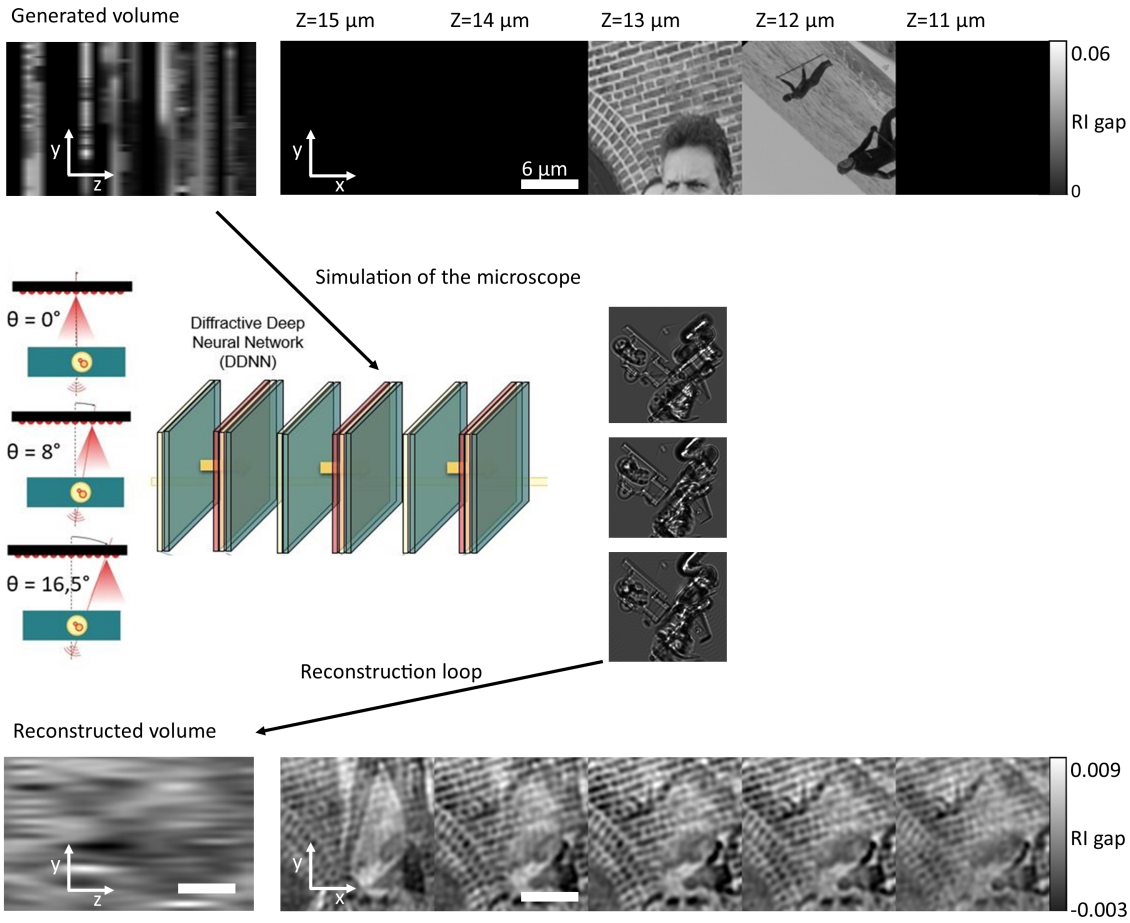


Figure 4.11: Example of a volume generated from random images and its simulated measurements, used to perform the reconstruction of this volume. All scale bars are 6 µm wide.

4.2.2.2 Model development and training

Now that we have paired 3D volumes with and without artefacts, we can use them to train a neural network to go from the distorted volume to the correct one.

Since we have 3D data, a straightforward strategy would be to use a 3D network such as a modified 3D U-net [17]. However, it has been shown that a multichannel 2D network would train faster and give similar results, provided that the multichannel input includes a central XY slice with various Z around it for each channel [83]. The idea is to give the network some 3D information, but not the whole volume, just 3 or 5 slice. We decided to use the same architecture that pix2pix [54] uses for image translation tasks. The only

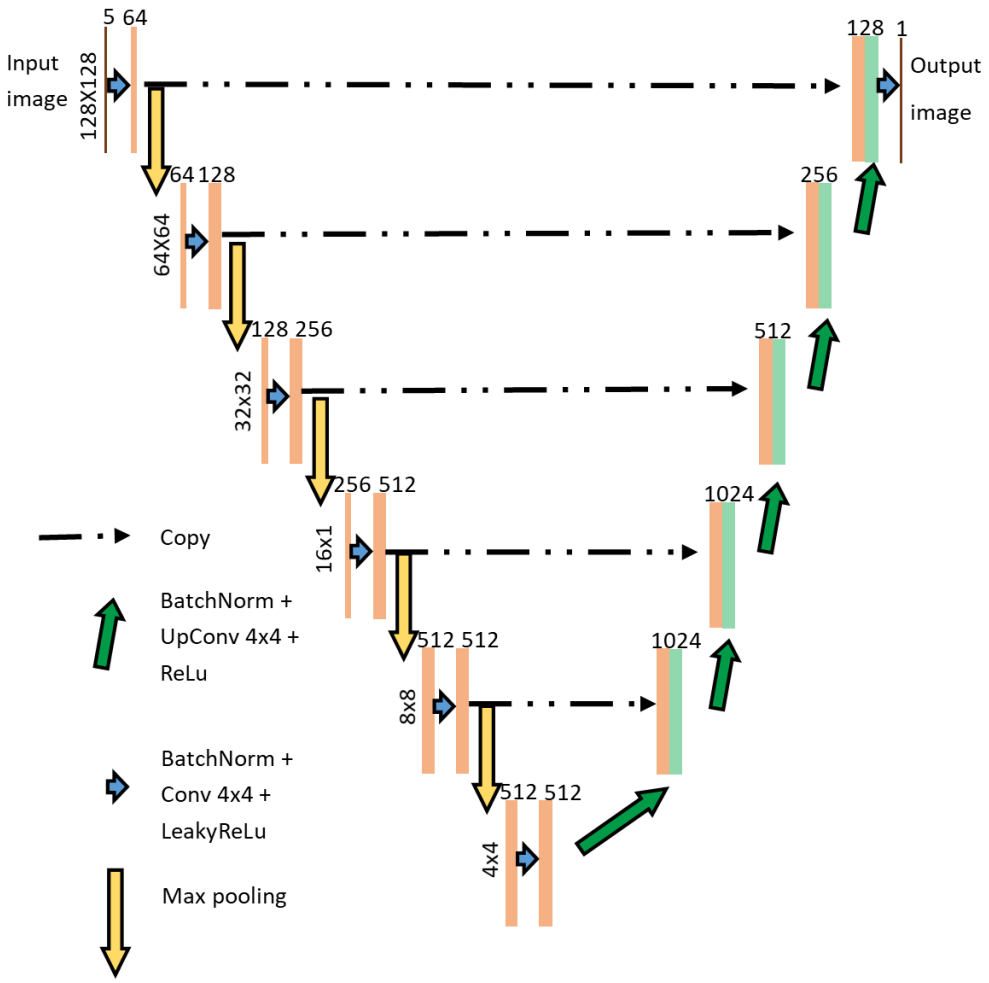


Figure 4.12: Architecture of the 2D modified U-net to take into account axial information with a 5 channel input.

modification was to use 5 input channels instead of the 3 standard channels for red, green, blue.

The main difference between the used architecture presented in Fig.4.12 and the original U-net is the use of modules of the form convolution, such as then BatchNorm [53] and the Leaky-ReLu. The BatchNorm layer helps the convergence and the leaky-ReLu helps the gradient to propagate during the backward passes.

This network was trained using Adam [64] algorithms with a learning rate of 1e-3 and a batch size of 20. One epoch correspond to 280 volumes. The data are selected randomly

before being fed to the neural network. All of this means that every 15 epochs, the network has seen every data once. The training and validation loss are visible on Fig.4.13. We are in a typical overfitting situation, where the training loss keep decreasing, while the validation loss remains constant and even increases after epochs 70. It is interesting to notice that the validation loss is always lower than the training loss during the first 30 epochs, meaning that the model performs better on unseen data. This can happen if the validation dataset is slightly easier, due to a lucky random draw of the images that created the volumes.

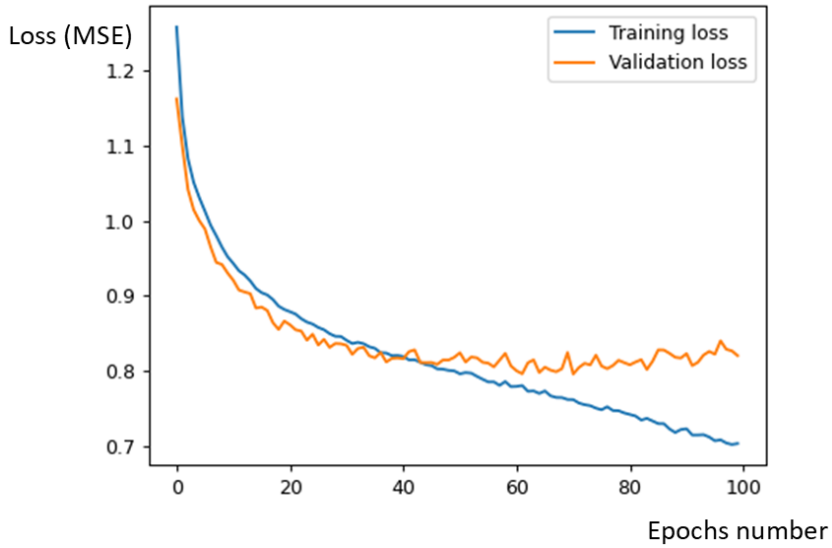


Figure 4.13: Loss evaluation during the training of our modified U-net. The training loss is the mean loss on the training dataset, the validation loss is the mean loss at the end of the epoch on the validation dataset.

An early stopping at epoch 61 was performed to keep the model that had the lowest validation loss. The performance on training and validation is shown on Fig.4.14.

The results on the training data are really encouraging. First every time that a image was supposed to be empty, the algorithm detected it as empty and returned a black (zero) image. Second, the elongation along the z direction disappeared, as excepted. Third, the image intensity (corresponding to the refractive index value) is much closer to the original image. For our microscopy setup, having a correct refractive index value is important because this parameter can be related to protein density or cell dry mass. Finally, the fine

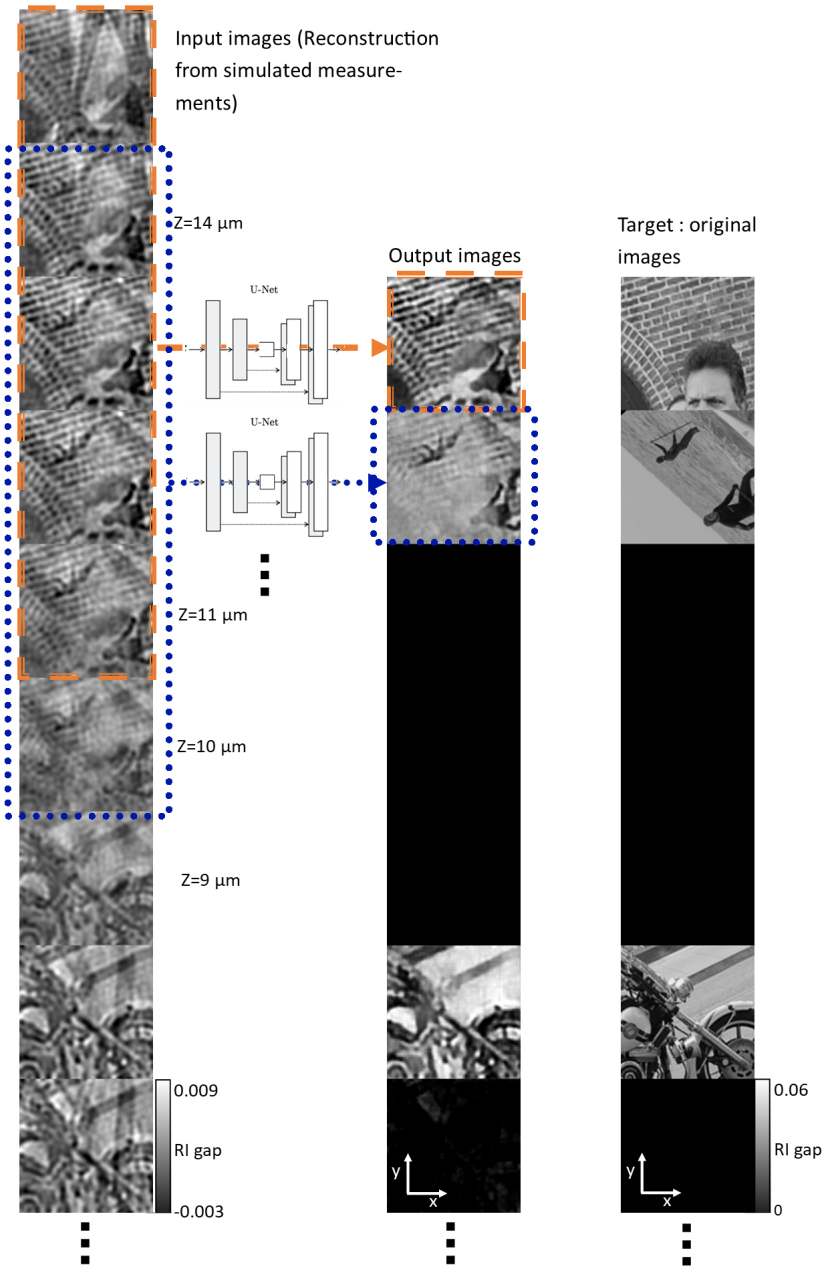


Figure 4.14: Visualisation of the output of the modified U-net on the training data created from random images.

details and high frequencies of the images are well preserved, as it can be seen in the bricks or on the wheel of the motorbike.

On the downside, we can see that the output from the neural network are not as clean

as the original images. Some details are blurry, and some important details tend to appear in the next or previous slice. (creating a 1µm or more elongation of object). For example the sky cannot be differentiated from the see in the surfer image.

The performance of any network on the training data is only relevant to make sure that the training went well and that the network is flexible enough to perform the desired task. To evaluate the performance on unseen data we need to use the validation dataset. An example of the output of the network on the validation dataset is shown on Fig.4.15.

The results are worst on this validation volume than on the training. Indeed, the validation dataset is much more complex to reconstruct, having a lot of different images next to each other, while on the training volume only 2 images were next to each other with a lot of empty refractive index examples before the third one. This complexity is visible on the reconstruction (the input data from the neural network), were all the images are mixed up and it is almost impossible for a human to guess what was the original image. Only the high contrast feature are found by the network, like the edge of picture are $z = 9\mu m$ or the reversed head at $z = 8\mu m$.

What is interesting on this validation volume is the correctness of the refractive index intensities, not going higher than 0.008 on the reconstruction and mapped them correctly to the good intensity range for each images (as high as 0.05). Another good point is the elongation correction, for $z = 7\mu m$ the output of the network is almost black, preventing any elongation of the head along the z direction. Concerning the elongation, the features present in one slice are not present in the neighbouring ones, i.e. every image is different from the one on top or below. This means that even if the low contrast features are lost, the objects or shapes detected are correctly placed along the z axis and they don't mix with each other. This fact is pretty stunning considering the fact that the axial resolution of our microscope with an objective of 0.4 NA and the red LED's with $\lambda = 0.625\ \mu m$ is $5\ \mu m$. This means that punctual objects are not supposed to be distinguished if they are not distant by at least 5 µm (theoretical resolution for the lateral direction). Here it looks like the modified U-net can separate high contrast features that are distant only by $1\mu m$.

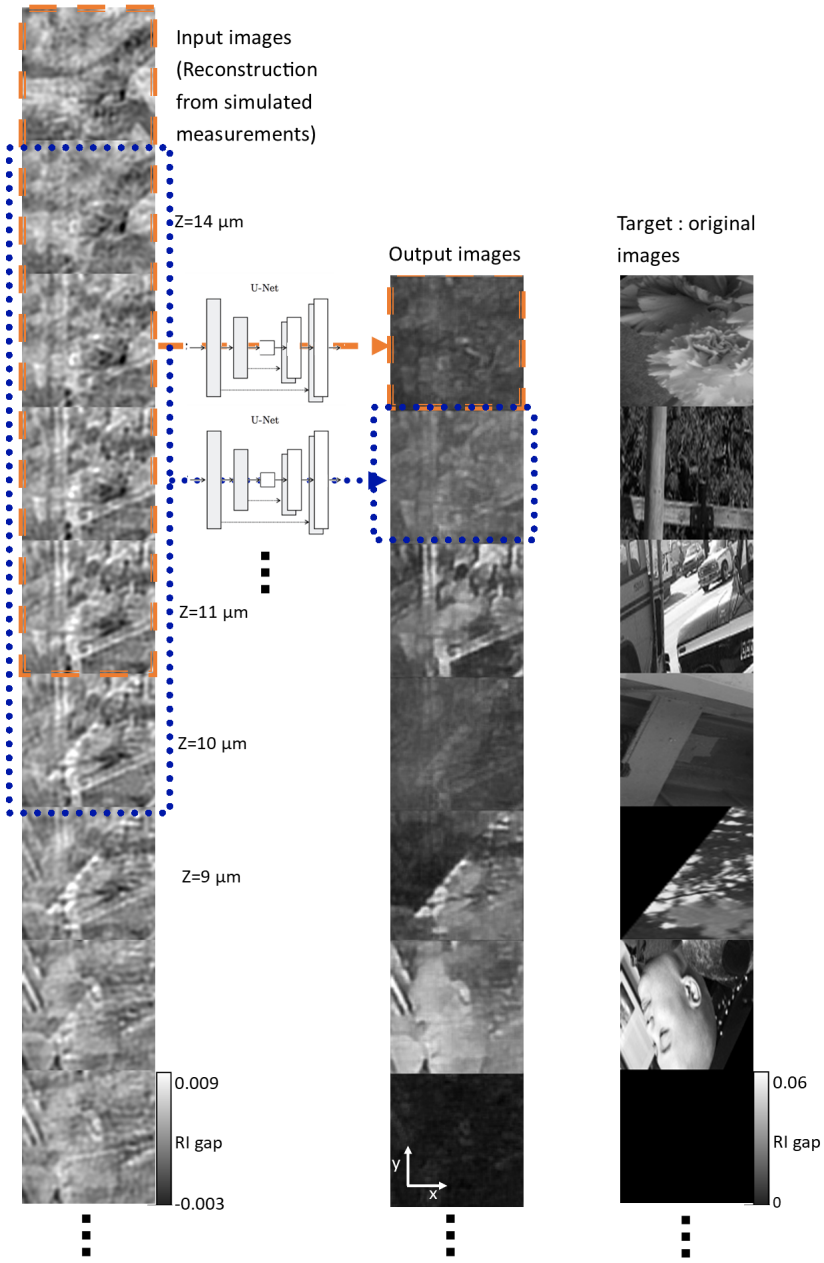


Figure 4.15: Visualisation of the output of the modified U-net on the validation data created from random images.

Concerning the fact that this modified U-net network cannot retrieve the low contrast objects or shapes when it is in the middle of many images, it might not be a relevant for biological objects. Indeed, the features inside a biological object are not changing that fast

along the z direction. See Fig.4.11 with YZ slice of the generated volume that is really far from a continuous slice inside a biological object because it contains a lot of white line. It looks closer to a barcode than an embryo. Concerning the refractive index gap, most features in which we are interested (lipid droplet, nuclei etc.) have a strong refractive index gap which is not the case here with a lot of low refractive index gap images.

4.2.2.3 Model testing on biological data

Now that our model is trained and that its performance is estimated on unseen simulation data, we can apply it to our real biological data.

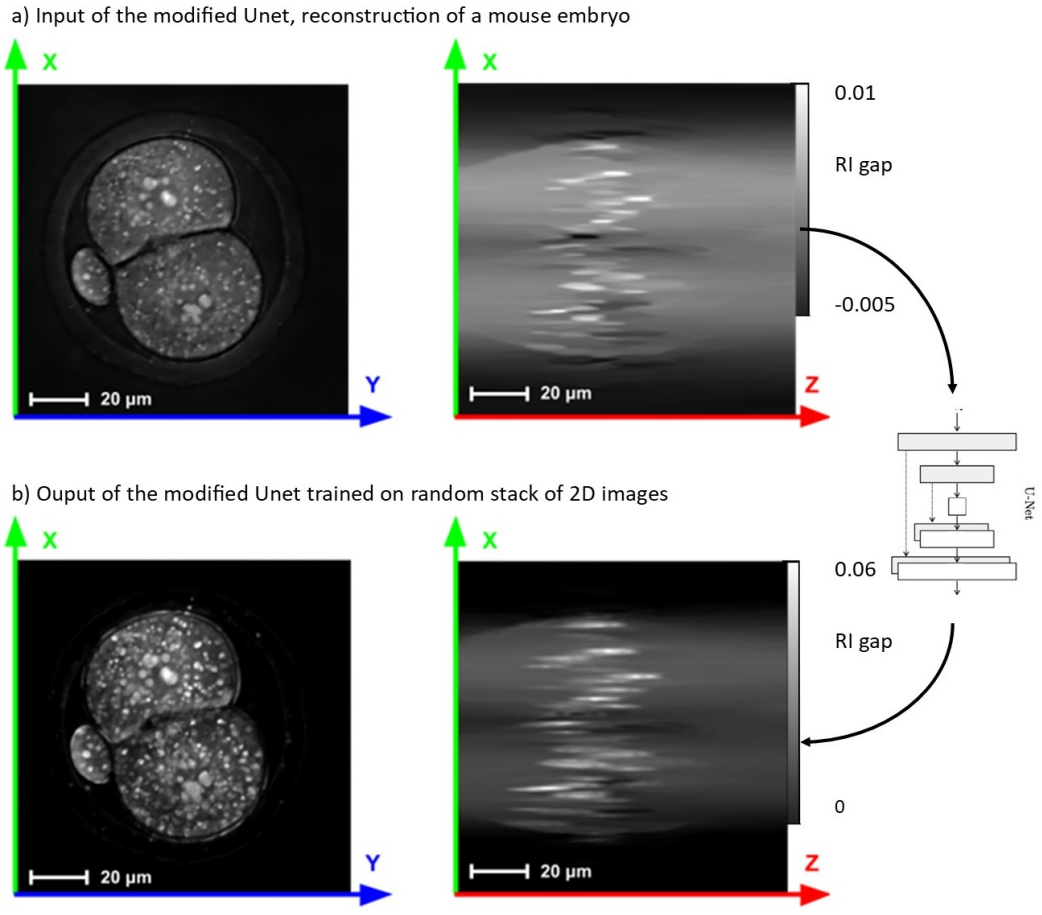


Figure 4.16: Performance of the modified U-net trained on random stack of images applied to a mouse embryo. Scale bars are $20\mu m$.

The chosen volume for this test is a 2 cell embryo from the time-lapse presented in Chapter 2. This volume is $1024 \times 1024 \times 128$ with the same spatial resolution than the $128 \times 128 \times 30$ volumes used for training (i.e. $0.128 \times 0.128 \times 1\mu m^3$). This embryo was reconstructed using the same algorithm than the simulated data. On Fig.4.16 we can see that the reconstructed refractive index range is the same for the embryo reconstructed with the BMP model and after passing the U-net architecture. The shape of 1024×1024 can be processed directly by the network; indeed the proposed U-net architecture is fully convolutional with max pooling factor of 2, therefore any power of 2 times 128 can pass forward.

The output of the neural network trained on random stack of image is presented in Fig.4.16. Since we do not have a reference image, only qualitative evaluation will be possible. The neural network had two main objectives: $i)$ removing the axial (along z axis) elongation, therefore increasing the refractive index value; $ii)$ filling the object refractive index to avoid the frequent "empty shell" aspect caused by the lack of phase information in the acquisition.

Concerning the elongation reduction, it is clearly visible that no reduction at all was performed by the neural network. Indeed, when we observe the nuclei or the liquid droplets that are almost spherical inside a mouse embryo, they appear with a strong elongation along the Z axis before and after the neural network. If we look closely at each droplet, we can see that its center is brighter. It might be explained by the fact that the neural network was trained on stacked images, and therefore was used to find in which slice a shape was originally from. But it never had to retrieve an object that was originally on multiple slices like a sphere. This might be the neural network answer to a sphere that he has never seen before. Even though the inner object of this embryo kept their elongated shape, their refractive was increased significantly, most likely leading to an overestimated refractive index.

Concerning the filling of object that only have their edges reconstructed, it is hard to tell whether the cells from this embryo are more compact before or after the neural network,

especially if we keep in mind the difference of color-map between the two images.

One may ask: why did it work that bad on real biological data ? A First answer would be the strong difference between the training data and the biological data. The network has never been confronted to any kind of continuous 3D shape along the z direction. Even though the refractive index ranges are similar over different layers, it appears that the network cannot still recognize continuous 3D shapes as single objects.

This answer is not fully satisfactory, because Matlock and Tian [83] managed to use their network trained on stack of images of various biological sample. Maybe we missed a crucial step in the data generation or in the neural network training, but the simplest answer to this question is that our biological and technical scenarios are more complex than what they were facing. In their case of intensity diffraction tomography the reconstructions were not satisfactory because of the use of a linear model. In our case, we are processing measurements with information content much lower as the incident wave phase is not recorded.

To find a solution to the elongation of reconstruction, we decided to change the object training to biological-like objects, with a particular focus on embryo-like objects. Here-below we describe the new methodology.

4.2.3 Artefact correction from simulated 3D biological volumes.

A large amount of 3D biological volumes are available, but they are mostly MRI images and CT scans. These imaging technique produce 3D volume always on grayscale and can easily be converted to a refractive index gap. The main issue with this idea is that most of the available dataset are very redundant in term of shape and contrast. Indeed a human lung will always look quite similar to an other human lung, and even when mixing all the body parts available (heart, liver, brain, knee etc.) the volumes were all looking quite similar. With heavy data augmentation it might have been to simulate the deformation from our microscope and train a network on it, but it seemed not diverse enough.

In microscopy, such 3D dataset exist but the number of cell imaged is still low [127]

or focused on 2D sample [20]. More importantly, there is no commercial reference microscope for refractive index imaging and all of the recent advances in optical diffraction tomography only include a low number of samples.

Considering the lack of synthetic 3D data available, we concluded that 3D synthetic data have to be generated.

4.2.3.1 Synthetic embryo generation

In the previous data generation we decided to promote diversity over biological-like features. This time, we will try to match as closely as possible the biological features while maintaining a high diversity of shapes and contrasts. Since no available dataset is available with such criteria, we created it from scratch.

From acquired images we observed that most of the features and shapes inside an embryo are some kind of 'blobs' or spheroid-like objects. These kinds of objects can be generated randomly and with a great amount of diversity.

The idea to generate a random spheroid-like object is to generate a sphere with variable radii in 3D. Perlin noise[93], originally designed to generate random map with hills and valley, was used to generate objects with various shapes by mapping this Perlin noise in 3D on the radius(see Fig.4.17).

To have an idea of the refractive index gap to use for the generated cell and nuclei we used Phasics (see section 3.2.5.4 for more detail) images and an hypothesis of sphericity of the object to go from a 2D quantitative phase image to the refractive index of object. Therefore the procedure to generate a synthetic embryo follows those three steps:

- Generate and place randomly between 10 and 20 synthetic cells. Each cell was created with a medium noise frequency and a low noise amplitude, as per Fig.4.17 a). Each cell is scaled randomly between 6 and 50 μm, with a refractive index gap randomly between 0.01 and 0.02.

- Inside each cell, add between 70 and 160 synthetic lipid droplets. Each lipid droplet

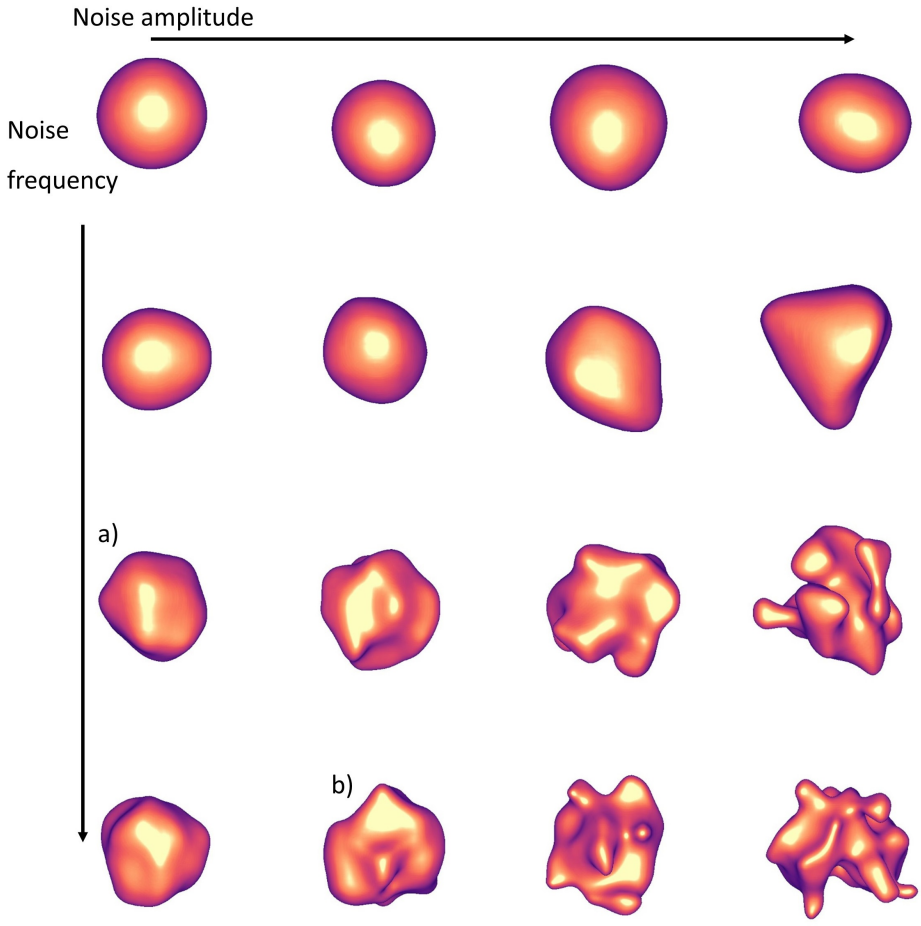


Figure 4.17: Evolution of the shape of a spheroid-like object generated with a random Perlin noise on the radius of a 3D sphere. a) was used to create the fake cell for the synthetic embryo. b) was used to create the fake lipid droplet for the synthetic embryo

is created with a high noise frequency and medium noise amplitude, as in Fig.4.17 b). Each lipid droplet is scaled randomly between 0.2 and 25 m and has a refractive index between 0.005 and 0.015 higher than its cell refractive index.

- Finally, an envelope is added around all the cells, representing the zona pellucida found on nominal embryos. This enevelope was created with a medium noise frequency and a low noise amplitude, as in Fig.4.17 a). This volume is then scaled randomly between 80 and 90 μm and the 90 % inside volume are removed. The refractive index was varied between 0.003 and 0.007.

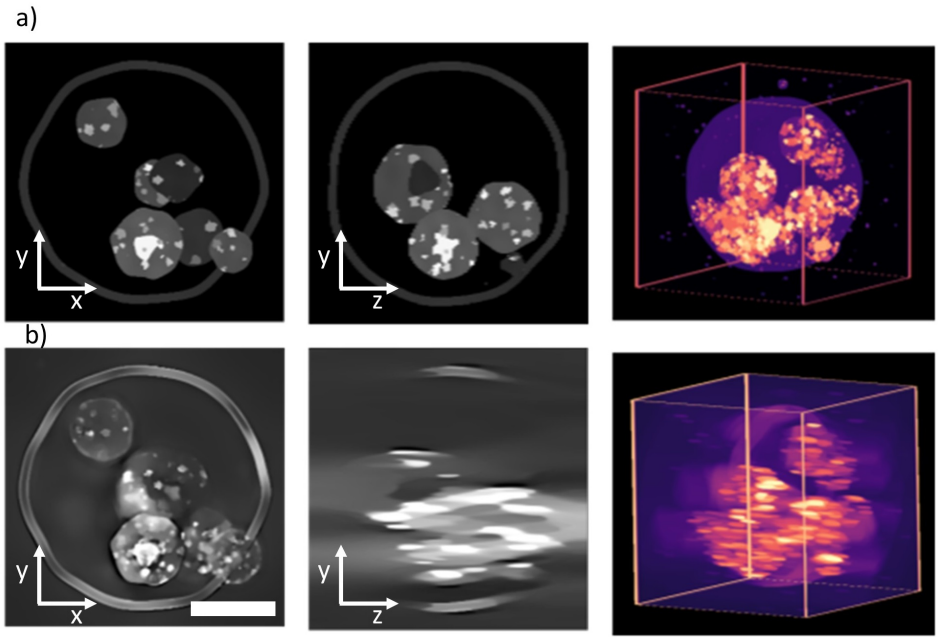


Figure 4.18: Visualisation of a synthetic embryo generated with the procedure from section 4.2.3.1. Scale bar is $40\ \mu m$ a) is the original generated volume. b) is the volume after simulated the optical system and doing the reconstruction from those generated measurement. The methodological procedure is the same as the one detailed in Section 4.2.2.1

A typical volume is presented in Fig.4.18, the same procedure than section 4.2.2.1 was used to produce the reconstruction from simulated measurement available in Fig 4.18 b, except for the volume size and discretization. Volumes of size $1024 \times 1024 \times 1024$ with a voxel size of $0.128 \times 0.128 \times 0.128 \mu m^3$ were used to simulated the microscope acquisition. From those acquisition, volumes of size $1024 \times 1024 \times 128$ with a voxel size of $0.128 \times 0.128 \times 1 \mu m^3$ were reconstructed.

It took 1 week using an un-optimised generation code and an optimized reconstruction code to generate 105 reference volume pairs. These volume pairs were used as training data for the neural network in the next section.

4.2.3.2 Neural network training and inference

The same modified U-net from section 4.2.2.2 was trained using Adam [64] algorithm with a learning rate of 1e-3. Random crops of $128 \times 128 \times 5$ were used for training with a batch

size of 20. One epoch correspond to 280 mini batch. Statistically, every 30 epochs, the network has seen every data once from the 105 $1024 \times 1024 \times 128$ volumes. The training and validation loss are visible on Fig.4.19. As usual, elongations over the z axis are clearly visible after the 3D reconstruction.

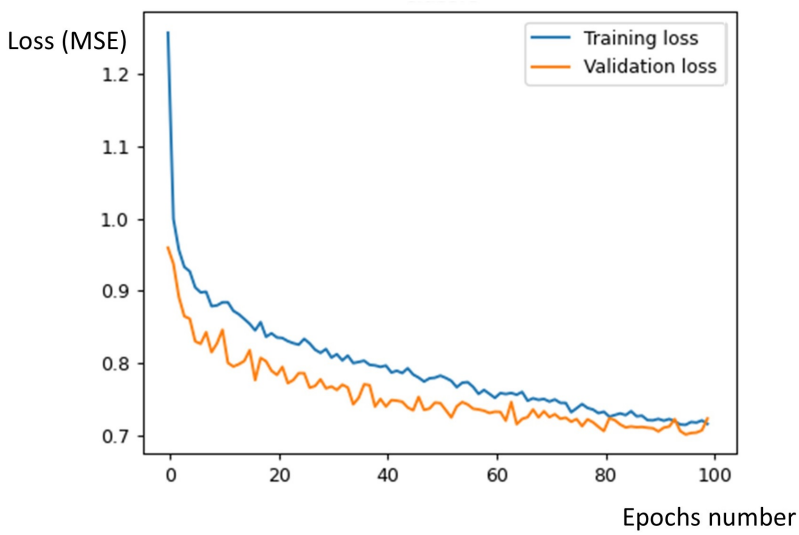


Figure 4.19: Loss evaluation during the training of our modified U-net on random volumes generated with perlin noise. The training loss is the mean loss on the training dataset, the validation loss is the mean loss at the end of the epoch on the validation dataset.

As observed in Fig.4.19, overfitting was not an issue on 100 epoch training. The convergence was well achieved, as the validation loss tend to a plateau.

The performance of this trained network on a validation volume is shown on Fig.4.20. The first thing to notice after the output of the modified U-net is how clean and continuous the xz cross-section is. This is impressive since each xy slice is produced separately and therefore we have no guaranty that the object generated in 3D will be continuous and without any noise like an object that appears only on one slice and disappears the next slice (this kind of flickering is a known issue with deep learning and temporal predictions [70]).

When looking at arrow number 1 on Fig.4.20 b), we can see that this synthetic cell with 3 synthetic droplets inside was suffering from heavy elongation along the z axis. Only

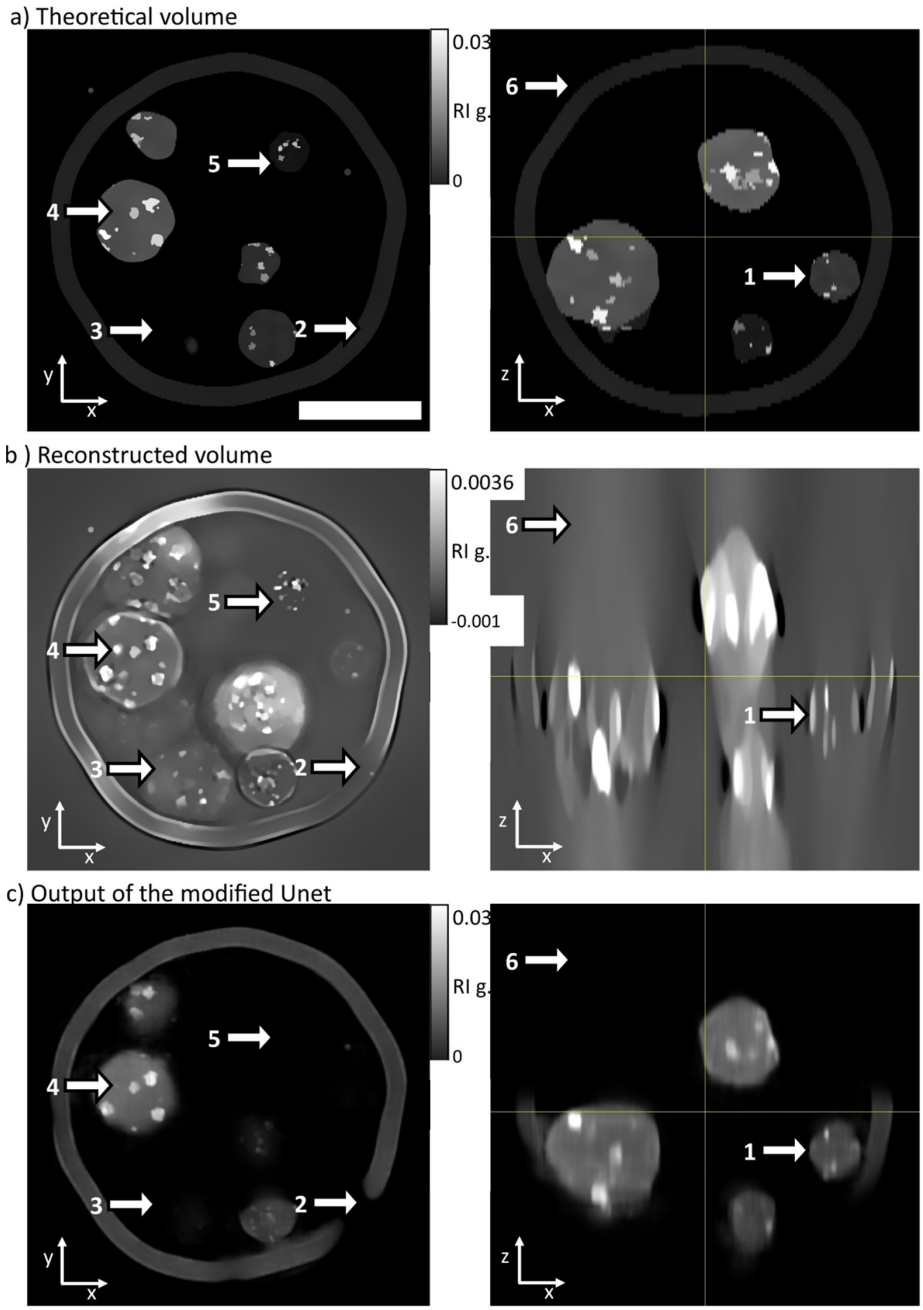


Figure 4.20: Qualitative evaluation of the modified U-net on the central (?) slice of a validation volume generated with Perlin noise. Scale bar is 40 µm.

the x and y edges of this cell were reconstructed, without clear boundary perceptible on the z axis. A human not familiarized with this kind of image would not have understood the true nature of this object. Only an expert used to work with embryos could have

guessed that it was an empty sphere with 3 dots inside. The neural network has clearly reached this level of comprehension, as it was able to find back the true xyz boundaries of the cell and to position correctly the inside synthetic droplets (Fig.4.20 c) arrow 1). The reduction of the elongation also benefits the xy cross section: as an example, the arrow 4 points at an inner structure that is not present in the theoretical volume and has appeared after the reconstructions. This structure exists in different layers of the volume, and only its elongation is visible in the studied slice. Our model truly identified this element and 5 others as not present in the original slice and correctly removed them, going back to 5 individual droplets, same than the reference volume in Fig.4.20 a) arrow 4. Inner structures can be removed while the cell remains, but entire cells can be removed too if they are not supposed to be on the slice, like the ones pointed by the arrow 3 or 5.

On the negative points of this model the removal of too many information is certainly the prominent one, like the cell pointed out with the arrow 5. This cell was partially present on the reconstructed volume, and should be present on the slice according to the theoretical volume (see the arrow 5 on Fig.4.20 c). The same bug happened for the bottom part of the ring around the cells (arrow 2). The outer membrane was not completely restored and only a partial ring was reconstructed. Considering the amount of information available in the reconstructed volume, it seems reassuring that the full outer membrane was not created everywhere (on the xz slice for example) because it would be too much guessing and not base on available information.

Even if most parts were correctly retrieved, some elements are missing. It is interesting to note that no element was hallucinated. Every object present in the output of the neural network truly exists in the theoretical volume.

4.2.3.3 Neural network evaluation on biological data

The results on simulated data are satisfactory enough to try the application of the proposed architecture on images of an embryo acquired with our setup.

The sample used for this test is the same one used in Section 4.2.2.3, a two cell embryo

from the time-lapse presented in Chapter 2. This volume is $1024 \times 1024 \times 128$ with the same spatial resolution than the volumes used for training. This embryo was reconstructed using the same algorithm than the simulated data.

On Fig.4.21 we can see the behaviour of the architecture on the mouse embryo. As expected, the cells are now filled with a higher refractive index than the reconstructed one. The z resolution of the inner lipid droplets is significantly improved after the application of the proposed U-net architecture, i.e. the droplets are a lot less elongated. Moreover, some lipid droplets that were indistinguishable on the standard reconstruction are now clearly separated. This network improves the images under different aspects: the elongation from inner structure as well as the one of the big cell are corrected, the empty shell effect on big objects is removed with a coherent refractive index filling, and finally the refractive index range is corrected to a higher value. This functioning network will be named "Uxy" for the remain part of this document.

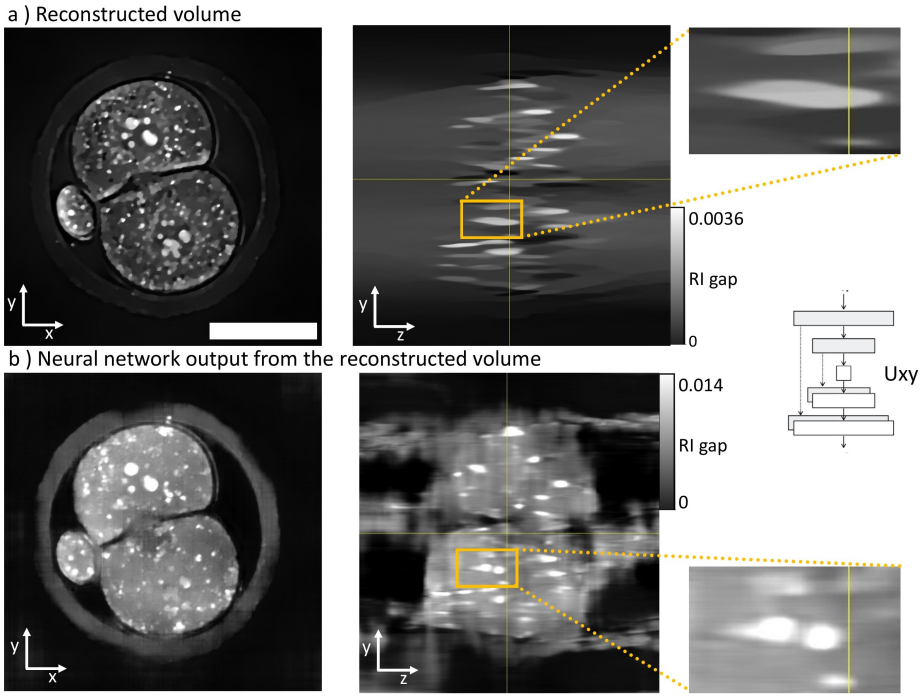


Figure 4.21: Qualitative evaluation of the modified U-net trained on synthetic volumes with Perlin noise. We named this model Uxy. The volume is tested on 2 cells mouse embryo. Scale bar is 40 µm.

The major drawback of Uxy is the lateral hallucination of the shape when an empty space sourrounds the object. Indeed, the contour of the cells in the z direction is neither lean nor correctly localized. This might be the limit of using 2D deep learning model to deal with a 3D problem. With more information along the z direction it looks trivial to say that nothing should be predicted at the top and bottom of the reconstructed volume. A last experiment that can be performed with 2D model but with more information along the z direction would be to make it predict yz slice or xz slice.

4.2.3.4 Training of a second network specialized along the axial direction

To understand the effects of an architecture that has more information on the axial direction (using yz or xz slice), a further training was performed using the same 3D training data already presented in 4.2.3.1.

The same neural network architecture was used, except for the first layer where mono channel 2D images were used instead of the 5 channel 2D images used in Uxy (see section 4.2.2.2 for more detail). The data-loader was of course modified to present randomly xz or yz slice. The amount of training data is doubled since we can pass each pixel twice in each epochs (inside a xz slice or inside a yz slice).

This model will be named Uz and can applied to all the xz slices of a reconstructed volume as well as all the yz slices. These two corrected volumes are presented in Fig.4.22. The performance of this network (Uz) is poor compared to Uxy. Indeed, the intracellular lipid droplets are less separated and a flickering effect appeared. This is caused by the fact that each xz or yz slice is generated separately from the others. If the neural network hesitates, it is common to have one slice white then black then white again etc. This creates the horizontal or vertical stripes. This issue was not present within Uxy model because each slice was predicted using multiple sequential slice, this means that at each prediction only a fifth of the input data was changing instead of 100% for the Uz model. A good idea to reduce the flickering effect on this model would be to change all its training loop and architecture and use multiple slices as input. This work was not done because even when

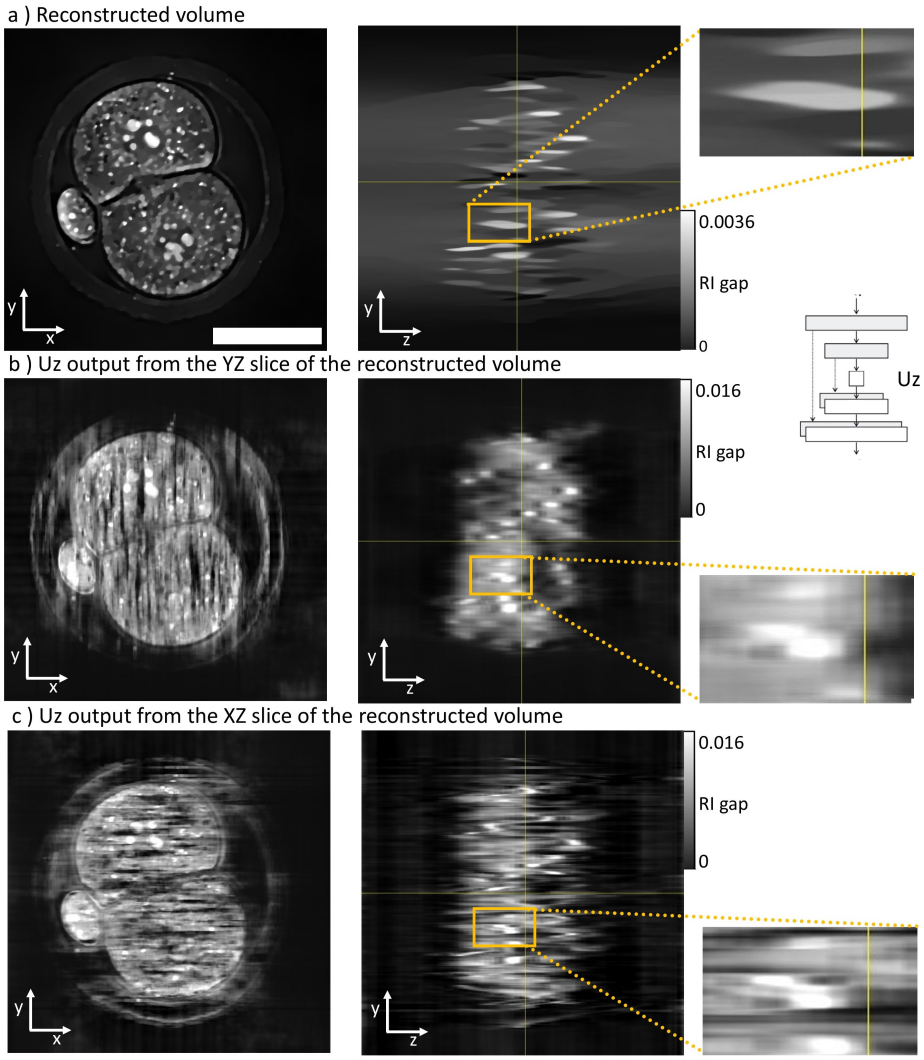


Figure 4.22: Qualitative evaluation of the Uz model trained on xz and yz slice from volumes generated with Perlin noise. The volume is tested on 2 cells mouse embryo. Scale bar is 40 µm.

looking at individual predicted slice like the yz slice from Fig.4.22 b), the predictions were less accurate than the one from Uxy.

The work done on this second network was important to give us hints on how to solve the drawbacks of Uxy and obtain meaningful 3D reconstructions. Indeed, the model Uz never hallucinates objects outside of the main target. An idea would be to use the output of the Uz as an envelope of the refractive index generated by Uxy.

4.2.3.5 *Ensembling of two neural networks: one specialised on the axial information and one specialized in the radial information*

The output of the network Uz trained with xz and yz slice is considered too noisy and presents undesired dark stripes. Some operations will be required to obtain a clean and smooth envelope from this neural network. The reconstructed embryo will be E_r. To have a more homogeneous volume we can sum up the volumes generated with Uz, $E_{sum} = Uz(E_r, XZ) + Uz(E_r, YZ)$. $Uz(x, A)$ is the results of the trained network Uz on the input data x, where the data are sliced in the plan A. A can be XZ or YZ.

E_{sum} will then be blurred with a Gaussian filter of size 20 and a threshold value of a fifth of the maximum value, create E_{bin}. It creates a finite support for the object that was inspired by [67].

$$E_{bin} = \begin{cases} 0, & \text{if } E_{sum} \circledast \sigma_{20} < \frac{1}{5} max(E_{sum}) \\ 1, & \text{otherwise} \end{cases} \tag{4.1}$$

This binary envelope will finally be blurred again with a Gaussian filter of size 3 to extend it slightly outside of its primary range, $E_{env} = E_{bin} \circledast \sigma_3$

The final output volume of our method will be the product between the enveloppe E_{env} created from Uz and the output of the network Uxy.

$$E_{final} = E_{env}.Uxy(E_r)$$

E_{final} is the artefact corrected volume on the bottom of the summary Fig.4.23 of our neural network ensembling method.

Using the method described above and summarized in Fig.4.23, the two cell embryo reconstructed volumes were corrected to reduce the axial elongation, fill constant object and correct the refractive index value.

An extensive comparison is performed in Fig.4.24. The elongation reduction that Uz could perform on its own is still present in the ensemble, allowing small inner structures

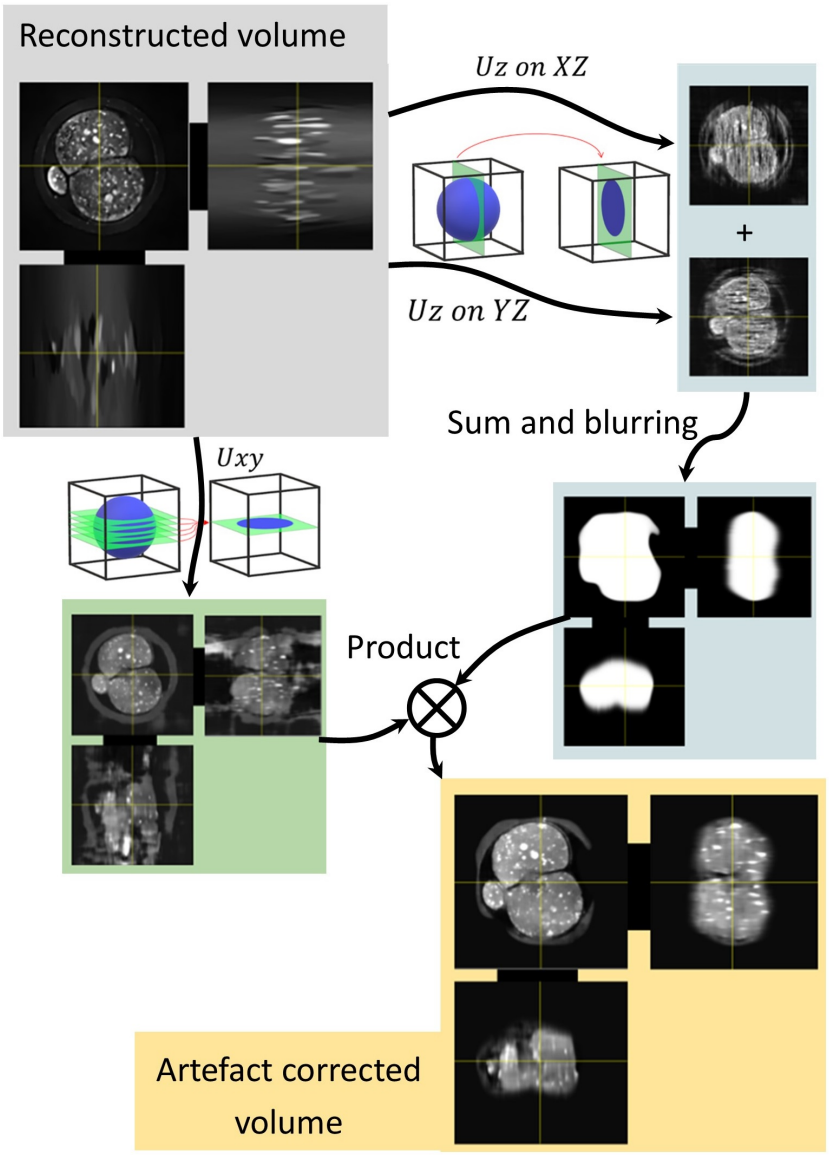


Figure 4.23: Summary of our ensembling method using two neural network, Uz specialized in the axial properties and Uxy specialized in the radial information.

to appear spherical in 3D, like the one pointed by the arrow 1. The 3D view allows the user to keep track of the particles while rotating the object. This is not possible with the non-corrected embryo, where every object is so elongated that they overlap each other. This elongation correction also separates object that appeared as one in the raw reconstructions, like the one pointed by arrow 2.

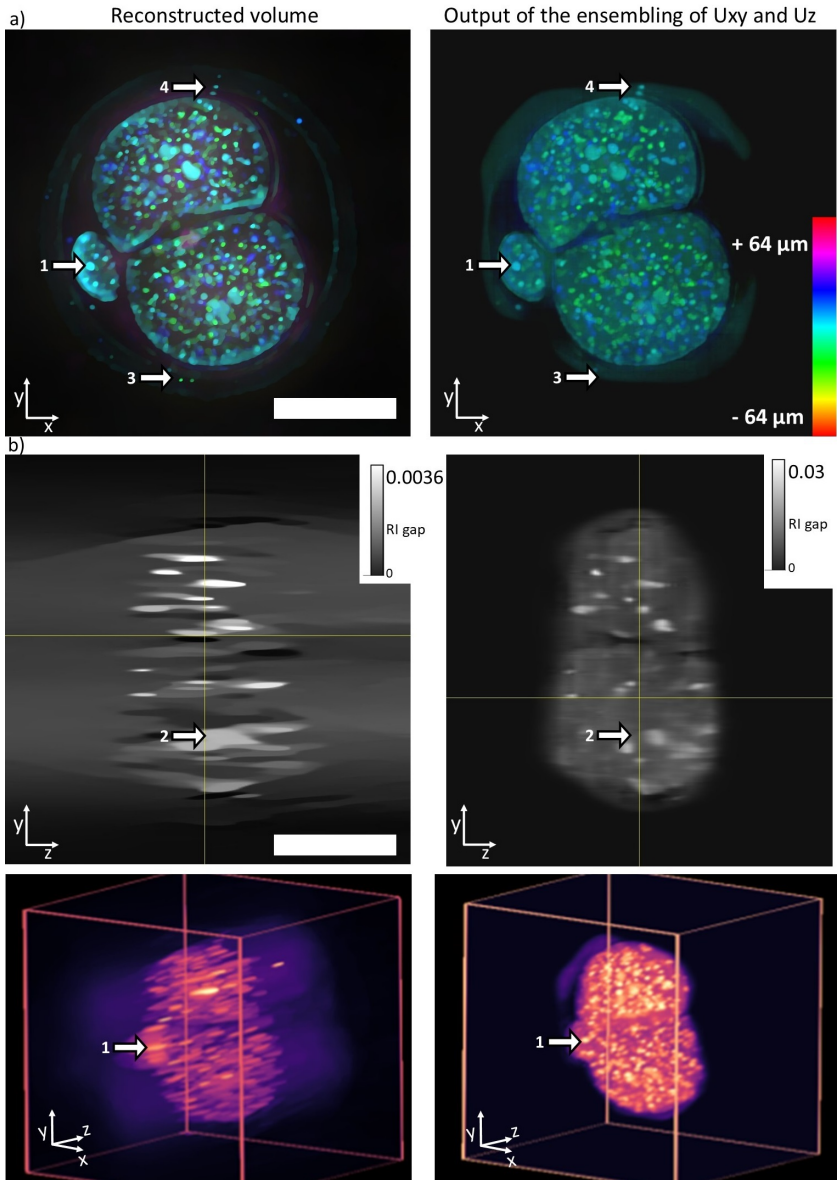


Figure 4.24: Comparison between the reconstructed volume and the output of the ensembling of the trained neural network Uz and Uxy. a) is a maximum intensity projection along the z axis with a color code for the z position. b) is a YZ slice where the elongation reduction of the neural network is clearly visible. c) is a 3D rendering. Scale bar is 40 µm wide

The contour of each cell is perfectly defined on the output of the neural network and no hallucinated object is present in the volume, which is a great improvement from the previous method in Section 4.2.3.3 with only Uz (see Fig **??** for the previous hallucinated

object).

On the drawback of this ensembling method, some small objects outside of the embryo where lost, like the one pointed by the arrow 3. This might be because the second network was used to compute the envelope of compact object and not to recover single small objects. To include small object outside of the main volumes could be an improvement direction considering that some small object on the side were correctly retrieved, arrow 4.

We can also state that the zona pellucida around the embryo that was barely visible on the reconstruction was only kept visible on half of the embryo on Fig.4.24 a).

We are satisfied by the result of this deep learning method for artefact correction on this 2 cell embryo. It is now time to address the generalisation question: will this method works for embryo at other stage of development ? And will this methods works for other biological object than embryo ? Each question will be answer in its own subsection.

4.2.3.6 Generalisation to other embryo stage

We saw in Chapter 2 that a complete time lapse of a mouse embryo was reconstructed in 3D. For each volume, the ensembling method of section 4.2.3.5 was used. Since the data were obtained with good experimental conditions, i.e. without vibration nor noticeable change in focusing distance, the neural network generalised well to embryo with only one cell, or multiple cells like the 4 cells embryo presented in Fig.4.25.

The same characteristics of this method were also present at every stage of the embryo development, as the elongation reduction or the good enveloping detection of each cell. however, the same drawbacks were also present with the issue of the outer membrane, the zona pellucida, that appears and disappears randomly.

It is important to remind us that these neural networks were trained only with simulated data and has never been confronted to real acquisitions before. The fact that they perform well and consistently with embryos at various stage of development is interesting, it showcases the generalisation capabilities of modern deep learning architecture in computer vision [145].

The choice of the training data is of paramount importance, and the main reason why the generalisation from simulated to real biological data was possible is because the generated data were carefully handcraft to mimic real data.

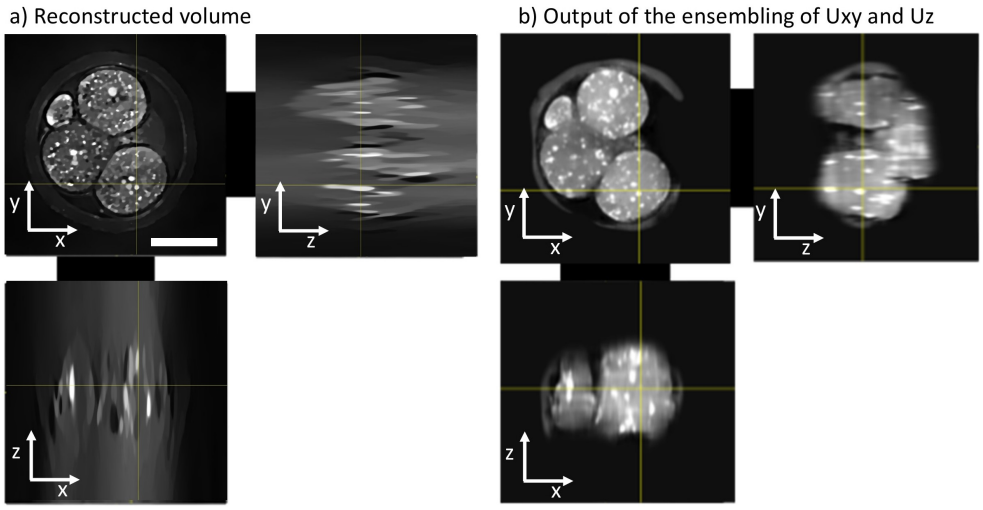


Figure 4.25: Comparison on a 4 cell mouse embryo between the reconstructed volume and the output of the ensembling of the trained neural network Uz and Uxy. a) is the reconstructed refractive index volume with 3 cross sections. b) is the 3 cross sections of the output of the ensembling of 2 neural networks specialised in radial and axial information. Scale bar is 40 µm wide

Indeed, with the volumes generated from random images almost no generalisation was observed. Without changing the training data is it possible for our algorithm to generalise to more compact and most complex living organism ?

4.2.3.7 Generalisation to denser and more complex biological sample

A good sample to try our ensembling model is the dense liver organoid from Chapter 2, Section 2.7.6. The oragnoid chosen was the one after 60 minutes of timelapse. Its reconstruction and the output of the neural networks is visible on Fig.4.26.

This object is not too far from an embryo because both the refractive range as well as the global size of the object are similar to the embryo . Moreover, there are some inner structures that produce elongated dots in the same manner than the embryo. However, the cells that compose the organoid are more entangled with each other that inside an embryo

and the hole that is present in the left part of the xy slice of the reconstructed volume is also new. .

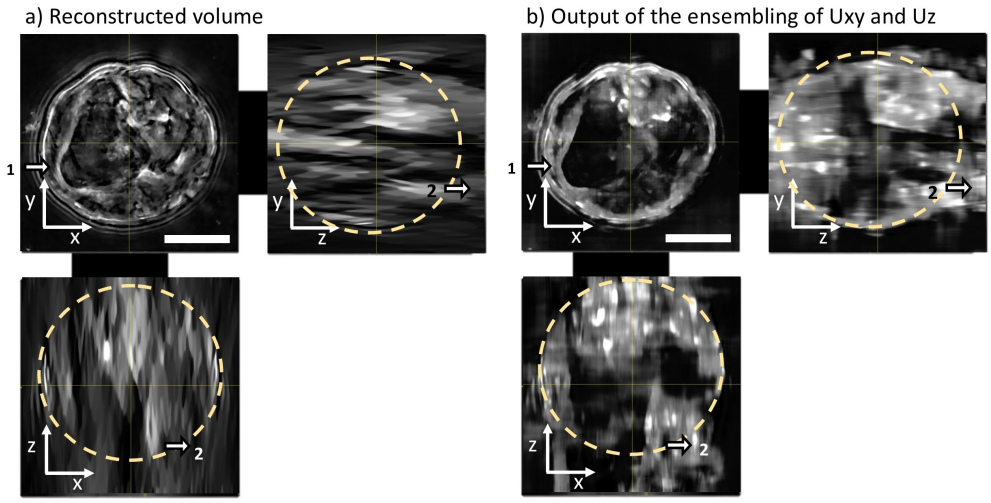


Figure 4.26: Comparison on a liver organoid between the reconstructed volume and the output of the ensembling of the trained neural network Uz and Uxy. a) is the reconstructed refractive index volume with 3 cross sections. b) is the 3 cross sections of the output of the ensembling of 2 neural networks specialised in radial and axial information. Scale bar is 40 µm wide.

Considering the refractive index range, our ensembling method seems to have increased the refractive index and also filled the cell regardless of their shape, like the one pointed from arrow 1 where only the edges were visible on Fig.4.26 a).

Concerning the elongation reduction, the few inner structures were correctly corrected to a better looking shape, but in the process, many non-existing inner structures were hallucinated by the network, pointed out by the arrows 2. Finally, considering the ability to find the correct envelope of the object and limit its spreading along the z axis, the networks performed poorly. Indeed, we expect the organoid to be rather spherical and no objects should appear outside of the dash circle. We would also expect a functional network to remove every refractive index shape from outside this circle, especially for very high of very low z. This is definitely not the case since the network predicts numerous features on the edges of the volume where it should be empty.

The generalization capabilities of our networks is not satisfactory toward non-embryo

biological object. Even with this non very challenging liver organoid, many drawbacks in terms of object hallucination or failure to understand the true shape in 3D makes this method non trustable to track the development of such organoid.

4.2.4 Conclusion on deep learning for artefact correction

Deep learning was used to tackle the main challenge of 3D refractive index imaging from intensity only images with a low NA objective and a wide field of view: the axial elongation of results. The state of the art in this field is still young and diverse, with deep learning blocks used at almost every points during the reconstruction loop. We decided to use deep learning models to correct the artefact present at the end of an iterative, regularized reconstruction with a multiple scattering model.

We have indeed demonstrated that it is possible to train a deep learning model to correct the major artefacts present in our TV regularized reconstruction with a BPM model using intensity only measurement from a low NA objective. Contrary to what was previously done with high quality optics but single scattering model, training on random stacks of images did not provide a good generalisation to 3D biological samples.

The training data need to be carefully created to match real refractive index distributions in terms of shape and RI range. Such dataset was generated with Perlin noise on random 3D spheroid. From those syntetic model, the measurements were simulated and reconstructed to create a 3D paired dataset of refractive index volumes. An ensemble of two 2D U-net models were used to correctly retrieve the refractive index in 3D.

The generalisation capacity of such network was good on embryos at various stage of development, which was the primary use case for this method. However the performance on liver organoid was unsatisfactory with hallucinated objects. That is why we think that the choice of the simulated training data is of paramount importance and the simulated objects need to be handcrafted with the support of an expert to be sure that the simulated volumes match real ones. A non-supervised approach might be envisaged to generalize the corrections to different samples and different acquisition parameters (e.g. different

magnifications or objective NAs), but this methodology is out of the scope of this thesis and was not pursued.

Chapter 5

Conclusion

5.1 GENERAL CONCLUSION

During this thesis we have explored the world of 3D microscopic refractive index imaging of biological sample. We choose to be as less intrusive as possible with the creation of a compact and cheap LED microscope. This prototype can image, in time-lapse, complex 3D biological sample developing inside the incubator. An efficient reconstruction algorithm that run on GPU was introduced to perform the 3D reconstruction with a multiple scattering model. It can run on a laptop as an oragnoids or an embryo develops for quasi-real time reconstruction. With simple hardware modification it can image up to a cubic millimeter volume and perform an acquisition every two minutes.

Chapter 2 presents five different light scattering model to solve the Helmholtz wave equation. Comparing to the Mie theory, the LS model is order of magnitude more accurate and more complex to compute than the other models introduced. It serves as a reference model to compare Born, Rytov, BPM and MBS (the first two are single scattering and the last two are multiple scattering). With our constraint of discretization size up to five times the wavelength, we choose to use the Beam Propagation Method (BPM) as our light scattering model for the rest of the thesis.

Chapter 3 introduces the inverse problem of optical diffraction tomography from intensity-

only measurements as an optimisation problem. It is solved through accelerate gradient descent with multiple priors such as Total variation, Sparsity and non-negativity. On simulation the importance of each prior is demonstrated. Then this reconstruction algorithm is used on real biological data with an extra layer of aberration correction for non-standard thickness container. By design our prototype was compact and compatible with biological culture conditions enabling us to image in time lapse the first week of a mouse embryo as well as the development of liver organoid.

Chapter 4 starts with a brief history of deep learning to have a better grasp on why it has revolutionised computer vision over the last decade. Deep learning strategies in optical diffraction tomography are reviewed and appear at every stage of the reconstruction algorithm. We choose to use a modified Unet to correct the main artefact that we were facing in Chapter 3 (*i*.e. axial elongation of object, and "edge only" object with a lack of low frequency). The final architecture ensembles two models trained on axial and radial information. Great results are obtained after a training only on simulated data. The training data was generated in the shape of fake multi cellular object and it generalizes well to real life embryo at various stage of development. The generalisation toward liver oragnoid is not good enough to be used routinely, with a lot of hallucinated objects.

5.2 DISCUSSION

Before diving into all the problems we are still facing and how to improve it with future works, let's remember that the computational system presented in this document is functional and robust. The preliminary results on embryo were foundation for further grant opportunity, the ANR project LIVE 3D_CNN (ANR-21-CE19-0020) and it is part of an European Union's Horizon 2020 research program, project REVEAL (grant agreement N° 101016726.).

The first drawback that is easy to improve is the imaging time of our setup. In order to address the imaging of living samples at a faster rate, one could use fewer and brighter

LEDs as in [76]. A ring of 8 powerful LEDs is placed on top of commercial inverted microscope, enabling them to image in 3D at a 10Hz rate. Reducing the number of image from 84 to 8 might lead to a higher noise or a slight loss of resolution.

The main drawback of our system is the strong axial elongation of object. This is not caused by flaw in our physical model or our reconstruction algorithm but is comes from a lack of information due to the missing cone problem well known in X-ray tomographic system. For this issue the state of the art method are learning tomography with sparse regularisation. To reduce the elongation higher NA optics needs to be used but the field of view will be reduce. Using deep learning method to "guess" the correct shape of object can provide stunning results, but the generalisation is never guarantied and the range of application of such supervised method is always limited. The lack of axial information for low NA diffractive tomographic microscope is still an open problem that will not be solved easily because of its inner lack of information.

Finally the focus needs to be discussed. Usually the imaged object is on focus at its center. When it is not the case because the object moved in 3D our resolution started to decreased with some artefacts appearing. The loss of resolution can also be observed for very thick objects where the most out of focus parts suffer from artefacts and the blurring effect. A study on the effect of defocus on diffractive tomographic microscope needs to be conducted to have a better understanding on this effect. For very thick object, one could try to image the same sample at multiple focuses in order to gain resolution everywhere inside the object.

Bibliography

[1] History of microscopy – timeline. https://www.sciencelearn.org.nz/resources/1692-history-of-microscopy-timeline. Accessed: 2022-01-28.

[2] Introductory confocal concepts. https://www.microscopyu.com/techniques/confocal/introductory-confocal-concepts. Accessed: 2021-01-05.

[3] Tomocube company website. https://www.tomocube.com. Accessed: 2020-01-21.

[4] Light sheet fluorescence microscopy. *Nature Reviews Methods Primers*, 1(1), November 2021.

[5] Rajat Acharya. Interaction of waves with medium. In *Satellite Signal Propagation, Impairments and Mitigation*, pages 57–86. Elsevier, 2017.

[6] S. Amini and S.M. Kirkup. Solution of helmholtz equation in the exterior domain by elementary boundary integral methods. *Journal of Computational Physics*, 118(2):208–221, 1995.

[7] M. Bernhardt, F. Wyrowski, and O. Bryngdahl. Coding and binarization in digital fresnel holography. *Optics Communications*, 77(1):4–8, June 1990.

[8] Dimitri P. Bertsekas. Incremental proximal methods for large scale convex optimization. *Mathematical Programming*, 129(2):163–195, June 2011.

[9] Eric Betzig, George H. Patterson, Rachid Sougrat, O. Wolf Lindwasser, Scott Olenych, Juan S. Bonifacino, Michael W. Davidson, Jennifer Lippincott-Schwartz, and Harald F. Hess. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, 313(5793):1642–1645, sep 2006.

[10] S Brenner. THE GENETICS OF CAENORHABDITIS ELEGANS. *Genetics*, 77(1):71–94, May 1974.

[11] Rich Caruana, Steve Lawrence, and C. Giles. Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping. In T. Leen, T. Dietterich, and V. Tresp, editors, *Advances in Neural Information Processing Systems*, volume 13. MIT Press, 2000.

[12] Michael Chen, David Ren, Hsiou-Yuan Liu, Shwetadwip Chowdhury, and Laura Waller. Multi-layer born multiple-scattering model for 3d phase microscopy. *Optica*, 7(5):394–403, May 2020.

[13] Michael Chen, Lei Tian, and Laura Waller. 3d differential phase contrast microscopy. *Biomed. Opt. Express*, 7(10):3940–3950, Oct 2016.

[14] Gunho Choi, DongHun Ryu, YoungJu Jo, Young Seo Kim, Weisun Park, Hyun seok Min, and YongKeun Park. Cycle-consistent deep learning approach to coherent noise reduction in optical diffraction tomography. *Opt. Express*, 27(4):4927–4943, Feb 2019.

[15] Shwetadwip Chowdhury, Michael Chen, Regina Eckert, David Ren, Fan Wu, Nicole Repina, and Laura Waller. High-resolution 3d refractive index microscopy of multiple-scattering samples from intensity images. 6(9):1211.

[16] Shwetadwip Chowdhury, Michael Chen, Regina Eckert, David Ren, Fan Wu, Nicole Repina, and Laura Waller. High-resolution 3d refractive index microscopy of multiple-scattering samples from intensity images. *Optica*, 6(9):1211–1219, Sep 2019.

[17] Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, and Olaf Ronneberger. 3d u-net: Learning dense volumetric segmentation from sparse annotation. *CoRR*, abs/1606.06650, 2016.

[18] Dan C. Ciresan, Ueli Meier, and Jürgen Schmidhuber. Multi-column deep neural networks for image classification. *CoRR*, abs/1202.2745, 2012.

[19] Laurent Condat. Discrete total variation: New definition and minimization. *SIAM Journal on Imaging Sciences*, 10(3):1258–1290, 2017.

[20] Ryan Conrad and Kedar Narayan. Cem500k, a large-scale heterogeneous unlabeled cellular electron microscopy image dataset for deep learning. *eLife*, 10:e65894, apr 2021.

[21] Yann Cotte, Fatih Toy, Pascal Jourdain, Nicolas Pavillon, Daniel Boss, Pierre Magistretti, Pierre Marquet, and Christian Depeursinge. Marker-free phase nanoscopy. *Nature Photonics*, 7(2):113–117, jan 2013.

[22] Y. H. Dai and Y. Yuan. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on Optimization*, 10(1):177–182, 1999.

[23] P. Davidovits and M. D. Egger. Scanning laser microscope for biological investigations. *Appl. Opt.*, 10(7):1615–1619, Jul 1971.

[24] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[25] L. Denneulin, F. Momey, D. Brault, M. Debailleul, A. M. Taddese, N. Verrier, and O. Haeberlé. Gsure criterion for unsupervised regularized reconstruction in tomographic diffractive microscopy. *J. Opt. Soc. Am. A*, 39(2):A52–A61, Feb 2022.

[26] A. J. Devaney. Inverse-scattering theory within the rytov approximation. 6(8):374–376. Publisher: Optical Society of America.

[27] Alberto Diaspro, Giuseppe Chirico, and Maddalena Collini. Two-photon fluorescence excitation and related techniques in biological microscopy. *Quarterly Reviews of Biophysics*, 38(2):97–166, 2005.

[28] Kenneth W. Dunn and Pamela A. Young. Principles of multiphoton microscopy. *Nephron Experimental Nephrology*, 103(2):e33–e40, mar 2006.

[29] Kenneth W. Dunn and Pamela A. Young. Principles of multiphoton microscopy. *Nephron Experimental Nephrology*, 103(2):e33–e40, March 2006.

[30] M. A. Dupertuis, B. Acklin, and M. Proctor. Generalization of complex snell–descartes and fresnel laws. *Journal of the Optical Society of America A*, 11(3):1159, March 1994.

[31] Eleftherios N. Economou. *Green's Functions in Quantum Physics*. Springer Berlin Heidelberg, 2006.

[32] Florian Faucher, Clemens Kirisits, Michael Quellmalz, Otmar Scherzer, and Eric Setterqvist. Diffraction Tomography, Fourier Reconstruction, and Full Waveform Inversion. 31 pages, 21 figures, October 2021.

[33] M. Feit and J. Fleck. Light propagation in graded-index optical fibers. *Applied optics*, 17 24:3990–8, 1978.

[34] J. R. Fienup. Phase retrieval algorithms: a comparison. 21(15):2758–2769. Publisher: Optical Society of America.

[35] Reeves Fletcher and Colin M Reeves. Function minimization by conjugate gradients. *The computer journal*, 7(2):149–154, 1964.

[36] Lord Rayleigh F.R.S. Xxxi. investigations in optics, with special reference to the spectroscope. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 8(49):261–274, 1879.

[37] Kunihiko Fukushima. Cognitron: A self-organizing multilayered neural network. *Biological Cybernetics*, 20(3-4):121–136, 1975.

[38] Kunihiko Fukushima. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 36:193–202, 1980.

[39] Greg Gbur and Emil Wolf. Diffraction tomography without phase information. *Opt. Lett.*, 27(21):1890–1892, Nov 2002.

[40] Joseph W Goodman. Introduction to fourier optics. *Introduction to Fourier optics, 3rd ed., by JW Goodman. Englewood, CO: Roberts & Co. Publishers, 2005*, 1, 2005.

[41] Joseph W Goodman. Introduction to fourier optics. *Introduction to Fourier optics, 3rd ed., by JW Goodman. Englewood, CO: Roberts & Co. Publishers, 2005*, 1, 2005.

[42] Enrico Gratton, Nicholas P. Barry, Sabrina Beretta, and Anna Celli. Multiphoton fluorescence microscopy. *Methods*, 25(1):103–110, 2001.

[43] Timur E. Gureyev, Timothy J. Davis, Andrew Pogany, Sheridan C. Mayo, and Stephen W. Wilkins. Optical phase retrieval by use of first born- and rytov-type approximations. 43(12):2418–2430.

[44] O. Haeberlé, K. Belkebir, H. Giovaninni, and A. Sentenac. Tomographic diffractive microscopy: basics, techniques and perspectives. *Journal of Modern Optics*, 57(9):686–699, 2010.

[45] L. Herve, O. Cioni, P. Blandin, F. Navarro, M. Menneteau, T. Bordy, S. Morales, and C. Allier. Multispectral total-variation reconstruction applied to lens-free microscopy. *Biomed. Opt. Express*, 9(11):5828–5836, Nov 2018.

[46] Samuel Hess, Thanu Girirajan, and Michael Mason. Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal*, 91:4258–72, 01 2007.

[47] Magnus R Hestenes and Eduard Stiefel. Methods of conjugate gradients for solving. *Journal of research of the National Bureau of Standards*, 49(6):409, 1952.

[48] Ralf Hielscher and Michael Quellmalz. Optimal mollifiers for spherical deconvolution. *Inverse Problems*, 31(8):085001, jun 2015.

[49] Robert Hooke. *Micrographia: or some physiological descriptions of minute bodies made by magnifying glasses, with observations and inquiries thereupon*. Courier Corporation, 2003.

[50] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989.

[51] Roarke Horstmeyer, Jaebum Chung, Xiaoze Ou, Guoan Zheng, and Changhuei Yang. Diffraction tomography with fourier ptychography. *Optica*, 3(8):827, jul 2016.

[52] B.J. Inkson. 2 - scanning electron microscopy (sem) and transmission electron microscopy (tem) for materials characterization. In Gerhard Hübschen, Iris Altpeter, Ralf Tschuncky, and Hans-Georg Herrmann, editors, *Materials Characterization Using Nondestructive Evaluation (NDE) Methods*, pages 17–43. Woodhead Publishing, 2016.

[53] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *CoRR*, abs/1502.03167, 2015.

[54] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks, 2016.

[55] Micah H. Jenkins and Thomas K. Gaylord. Three-dimensional quantitative phase imaging via tomographic deconvolution phase microscopy. *Appl. Opt.*, 54(31):9213–9227, Nov 2015.

[56] Di Jin, Renjie Zhou, Zahid Yaqoob, and Peter T. C. So. Tomographic phase microscopy: principles and applications in bioimaging [invited]. *Journal of the Optical Society of America B*, 34(5):B64, apr 2017.

[57] YoungJu Jo, Sangjin Park, JaeHwang Jung, Jonghee Yoon, Hosung Joo, Min hyeok Kim, Suk-Jo Kang, Myung Chul Choi, Sang Yup Lee, and YongKeun Park. Holographic deep learning for rapid optical screening of anthrax spores. *Science Advances*, 3(8), August 2017.

[58] Ulugbek S. Kamilov, Ioannis N. Papadopoulos, Morteza H. Shoreh, Alexandre Goy, Cedric Vonesch, Michael Unser, and Demetri Psaltis. Learning approach to optical tomography. *Optica*, 2(6):517.

[59] Ulugbek S. Kamilov, Ioannis N. Papadopoulos, Morteza H. Shoreh, Alexandre Goy, Cedric Vonesch, Michael Unser, and Demetri Psaltis. Optical tomographic image reconstruction based on beam propagation and sparse regularization. *IEEE Transactions on Computational Imaging*, 2(1):59–70, 2016.

[60] Iksung Kang, Alexandre Goy, and George Barbastathis. Computational optical tomography using 3-d deep convolutional neural networks. *Optical Engineering*, 57(04):1, apr 2018.

[61] Iksung Kang, Alexandre Goy, and George Barbastathis. Limited-angle tomographic reconstruction of dense layered objects by dynamical machine learning, 2020.

[62] Myung K. Kim. Principles and techniques of digital holographic microscopy. *SPIE Reviews*, 1(1):018005, 2010.

[63] Tae In Kim, Buki Kwon, Jonghee Yoon, Ick-Joon Park, Gyeong Sook Bang, YongKeun Park, Yeon-Soo Seo, and Sung-Yool Choi. Antibacterial activities of graphene oxide–molybdenum disulfide nanocomposite films. *ACS Applied Materials*, 9(9):7908–7917, February 2017.

[64] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization, 2014.

[65] Clemens Kirisits, Michael Quellmalz, Monika Ritsch-Marte, Otmar Scherzer, Eric Setterqvist, and Gabriele Steidl. Fourier reconstruction for diffraction tomography of an object rotated into arbitrary orientations. *Inverse Problems*, 37(11):115002, oct 2021.

[66] Tobias Knopp, Stefan Kunis, and Daniel Potts. A note on the iterative MRI reconstruction from nonuniform space data. *International Journal of Biomedical Imaging*, 2007:1–9, 2007.

[67] Wojciech Krauze. Optical diffraction tomography with finite object support for the minimization of missing cone artifacts. *Biomed. Opt. Express*, 11(4):1919–1926, Apr 2020.

[68] Małgorzata Kujawińska, Wojciech Krauze, Arkadiusz Kus, Julianna Kostencka, Tomasz Kozacki, Björn Kemper, and Michał Dudek. Problems and solutions in 3-d analysis of phase biological objects by optical diffraction tomography. *International Journal of Optomechatronics*, 8(4):357–372, 2014.

[69] Knut Kvien. Validity of weak-scattering models in forward two-dimensional optical scattering. *Applied Optics*, 34(36):8447, December 1995.

[70] Wei-Sheng Lai, Jia-Bin Huang, Oliver Wang, Eli Shechtman, Ersin Yumer, and Ming-Hsuan Yang. Learning blind video temporal consistency. *CoRR*, abs/1808.00449, 2018.

[71] P Philippe Laissue, Rana A Alghamdi, Pavel Tomancak, Emmanuel G Reynaud, and Hari Shroff. Assessing phototoxicity in live fluorescence imaging. *Nature Methods*, 14(7):657–661, July 2017.

[72] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, may 2015.

[73] Peng Lee, Wanrong Gao, and Xianling Zhang. Performance of single-scattering model versus multiple-scattering model in the determination of optical properties of biological tissue with optical coherence tomography. *Applied Optics*, 49(18):3538, June 2010.

[74] Chen Li, Adele Moatti, Xuying Zhang, H. Troy Ghashghaei, and Alon Greenbaum. Deep learning-based autofocus method enhances image quality in light-sheet fluorescence microscopy. *Biomed. Opt. Express*, 12(8):5214–5226, Aug 2021.

[75] Jiaji Li, Qian Chen, Jiasong Sun, Jialin Zhang, Junyi Ding, and Chao Zuo. Three-dimensional tomographic microscopy technique with multi-frequency combination with partially coherent illuminations. *Biomed. Opt. Express*, 9(6):2526–2542, Jun 2018.

[76] Jiaji Li, Alex C. Matlock, Yunzhe Li, Qian Chen, Chao Zuo, and Lei Tian. High-speed in vitro intensity diffraction tomography. *Advanced Photonics*, 1(6):1 – 13, 2019.

[77] Qingxiang Li, Lifen Bai, Shifu Xue, and Luyun Chen. Autofocus system for microscope. *Optical Engineering*, 41(6):1289 – 1294, 2002.

[78] Joowon Lim, Ahmed B. Ayoub, Elizabeth E. Antoine, and Demetri Psaltis. High-fidelity optical diffraction tomography of multiple scattering samples. 8(1):82.

[79] Joowon Lim, Ahmed B. Ayoub, and Demetri Psaltis. Three-dimensional tomography of red blood cells using deep learning. *Advanced Photonics*, 2(02):1, mar 2020.

[80] JooWon Lim, KyeoReh Lee, Kyong Hwan Jin, Seungwoo Shin, SeoEun Lee, YongKeun Park, and Jong Chul Ye. Comparative study of iterative reconstruction algorithms for missing cone problems in optical diffraction tomography. *Opt. Express*, 23(13):16933–16948, Jun 2015.

[81] Tsung-Yi Lin, Michael Maire, Serge J. Belongie, Lubomir D. Bourdev, Ross B. Girshick, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. Microsoft COCO: common objects in context. *CoRR*, abs/1405.0312, 2014.

[82] Yifang Ma, Zhenyu Wang, Hong Yang, and Lin Yang. Artificial intelligence applications in the development of autonomous vehicles: a survey. *IEEE/CAA Journal of Automatica Sinica*, 7(2):315–329, March 2020.

[83] Alex Matlock and Lei Tian. Physical model simulator-trained neural network for computational 3d phase imaging of multiple-scattering samples. *arXiv*, 03 2021.

[84] Francesco Merola, Pasquale Memmolo, Lisa Miccio, Roberto Savoia, Martina Mugnano, Angelo Fontana, Giuliana D'Ippolito, Angela Sardo, Achille Iolascon, Antonella Gambale, and Pietro Ferraro. Tomographic flow cytometry by digital holography. *Light: Science & Applications*, 6(4):e16241–e16241, October 2016.

[85] Gustav Mie. Beiträge zur optik trüber medien, speziell kolloidaler metallösungen. *Annalen der Physik*, 330(3):377-445, Apr 1908.

[86] Kristina Monakhova, Joshua Yurtsever, Grace Kuo, Nick Antipa, Kyrollos Yanny, and Laura Waller. Learned reconstructions for practical mask-based lensless imaging. *Opt. Express*, 27(20):28075–28090, Sep 2019.

[87] Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. Deepstack: Expert-level artificial intelligence in no-limit poker. *Science*, 356, 01 2017.

[88] Paul Müller, Mirjam Schürmann, and Jochen Guck. The theory of diffraction tomography, 2016.

[89] Tan H. Nguyen, Mikhail E. Kandel, Marcello Rubessa, Matthew B. Wheeler, and Gabriel Popescu. Gradient light interference microscopy for 3d imaging of unlabeled specimens. *Nature Communications*, 8(1), August 2017.

[90] Xiaoze Ou, Guoan Zheng, and Changhuei Yang. Embedded pupil function recovery for fourier ptychographic microscopy. *Opt. Express*, 22(5):4960–4972, Mar 2014.

[91] Francesco Pampaloni, Bo-Jui Chang, and Ernst H. K. Stelzer. Light sheet-based fluorescence microscopy (LSFM) for the quantitative imaging of cells and tissues. *Cell and Tissue Research*, 360(1):129–141, mar 2015.

[92] V.Y. Panin, G.L. Zeng, and G.T. Gullberg. Total variation regulated em algorithm [spect reconstruction]. *IEEE Transactions on Nuclear Science*, 46(6):2202–2210, 1999.

[93] Ken Perlin. An image synthesizer. *ACM Siggraph Computer Graphics*, 19(3):287–296, 1985.

[94] 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, feb 2001.

[95] Mikael Persson, Dianna Bone, and Håkan Elmqvist. Three-dimensional total variation norm for spect reconstruction. *Nuclear Instruments & Methods in Physics Research Section A-accelerators Spectrometers Detectors and Associated Equipment - NUCL INSTRUM METH PHYS RES A*, 471:98–102, 09 2001.

[96] T. Pham, E. Soubies, A. Ayoub, J. Lim, D. Psaltis, and M. Unser. Three-dimensional optical diffraction tomography with lippmann-schwinger model. *IEEE Transactions on Computational Imaging*, 6:727–738, 2020.

[97] William Pierré, Lionel Hervé, Cédric Allier, Sophie Morales, Sergei Grudinin, Shwetadwip Chowdhury, Laura Waller, Christophe ARNOULT, Pierre RAY, and Magali Dhellemmes. Deep learning framework applied to optical diffraction tomography (ODT). In Thomas G. Brown, Tony Wilson, and Laura Waller, editors, *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XXVIII*. SPIE, March 2021.

[98] William Pierré, Lionel Hervé, Cédric Allier, Sophie Morales, Sergei Grudinin, Pierre RAY, Christophe Arnoult, and Magali Dhellemmes. 3D live cell imaging of whole organoids in time-lapse using intensity diffraction tomography. In Marc P. Georges, Gabriel Popescu, and Nicolas Verrier, editors, *Unconventional Optical Imaging III*, volume PC12136, page PC1213612. International Society for Optics and Photonics, SPIE, 2022.

[99] William Pierré, Lionel Hervé, Chiara Paviolo, Ondrej Mandula, Vincent Remondiere, Sophie Morales, Sergei Grudinin, Pierre F. Ray, Magali Dhellemmes, Christophe Arnoult, and Cédric Allier. 3d time-lapse imaging of a mouse embryo using intensity diffractiontomography embedded inside a deep learning framework. *Appl. Opt.*, 61(12):3337–3348, Apr 2022.

[100] Boris Teodorovich Polyak. The conjugate gradient method in extremal problems. *USSR Computational Mathematics and Mathematical Physics*, 9(4):94–112, 1969.

[101] Daniel Potts and Gabriele Steifdl. A new linogram algorithm for computerized tomography. *IMA Journal of Numerical Analysis*, 21, 07 2001.

[102] Malte Renz. Fluorescence microscopy-a historical and technical perspective. *Cytometry Part A*, 83(9):767–779, April 2013.

[103] Yair Rivenson, Tairan Liu, Zhensong Wei, Yibo Zhang, Kevin de Haan, and Aydogan Ozcan. PhaseStain: the digital staining of label-free quantitative phase microscopy images using deep learning. *Light: Science Applications*, 8(1), feb 2019.

[104] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

[105] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408, 1958.

[106] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[107] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[108] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, oct 1986.

[109] Michael J Rust, Mark Bates, and Xiaowei Zhuang. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature Methods*, 3(10):793–796, aug 2006.

[110] DongHun Ryu, YoungJu Jo, Jihyeong Yoo, Taean Chang, Daewoong Ahn, Young Seo Kim, Geon Kim, Hyun-Seok Min, and YongKeun Park. Deep learning-based optical field screening for robust optical diffraction tomography. *Scientific Reports*, 9(1), oct 2019.

[111] DongHun Ryu, Jinho Kim, Daejin Lim, Hyun-Seok Min, In Young Yoo, Duck Cho, and YongKeun Park. Label-free white blood cell classification using refractive index tomography and deep learning. *BME Frontiers*, 2021:1–9, jul 2021.

[112] Amirhossein Saba, Carlo Gigli, Ahmed B. Ayoub, and Demetri Psaltis. Physics-informed neural networks for diffraction tomography, 2022.

[113] Vijay Sankaran, Charles M. Weber, Fred Lakhani, and Jr. Kenneth W. Tobin. Inspection in semiconductor manufacturing, December 1999.

[114] Adam Santoro, David Raposo, David G Barrett, Mateusz Malinowski, Razvan Pascanu, Peter Battaglia, and Timothy Lillicrap. A simple neural network module for relational reasoning. In I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017.

[115] C.J.R. Sheppard and A. Choudhury. Image formation in the scanning microscope. *Optica Acta: International Journal of Optics*, 24(10):1051–1073, 1977.

[116] Seungwoo Shin, Doyeon Kim, Kyoohyun Kim, and YongKeun Park. Super-resolution three-dimensional fluorescence and optical diffraction tomography of live cells using structured illumination generated by a digital micromirror device. *Scientific Reports*, 8(1), jun 2018.

[117] Seungwoo Shin, Kyoohyun Kim, Jonghee Yoon, and YongKeun Park. Active illumination using a digital micromirror device for quantitative phase imaging. *Opt. Lett.*, 40(22):5407–5410, Nov 2015.

[118] Emil Y. Sidky, Chien-Min Kao, and Xiaochuan Pan. Accurate image reconstruction from few-views and limited-angle data in divergent-beam ct. 2009.

[119] Bertrand Simon, Matthieu Debailleul, Anne Beghin, Yves Tourneur, and Olivier Haeberlé. High-resolution tomographic diffractive microscopy of biological samples. *Journal of Biophotonics*, 3(7):462–467, March 2010.

[120] M. Slaney, A.C. Kak, and L.E. Larsen. Limitations of imaging with first-order diffraction tomography. *IEEE Transactions on Microwave Theory and Techniques*, 32(8):860–874, August 1984.

[121] Juan M. Soto, José A. Rodrigo, and Tatiana Alieva. Label-free quantitative 3d tomographic imaging for partially coherent light microscopy. *Opt. Express*, 25(14):15699–15712, Jul 2017.

[122] Ferreol Soulez. Gauging diffraction patterns: field of view and bandwidth estimation in lensless holography. *Appl. Opt.*, 60(10):B38–B48, Apr 2021.

[123] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014.

[124] Frederic Strobl, Alexander Schmitz, and Ernst H K Stelzer. Improving your four-dimensional image: traveling through a decade of light-sheet-based fluorescence microscopy research. *Nature Protocols*, 12(6):1103–1109, may 2017.

[125] David Strong and Tony Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19(6):S165–S187, nov 2003.

[126] Yongjin Sung, Wonshik Choi, Christopher Fang-Yen, Kamran Badizadegan, Ramachandra R. Dasari, and Michael S. Feld. Optical diffraction tomography for high resolution live cell imaging. *Opt. Express*, 17(1):266–277, Jan 2009.

[127] Ahmad P. Tafti, Andrew B. Kirkpatrick, Jessica D. Holz, Heather A. Owen, and Zeyun Yu. 3dsem: A 3d microscopy dataset. *Data in Brief*, 6:112–116, 2016.

[128] Waleed Tahir, Ulugbek S. Kamilov, and Lei Tian. Holographic particle localization under multiple scattering. *Advanced Photonics*, 1(03):1, June 2019.

[129] Lei Tian, Xiao Li, Kannan Ramchandran, and Laura Waller. Multiplexed coded illumination for fourier ptychography with an led array microscope. *Biomed. Opt. Express*, 5(7):2376–2389, Jul 2014.

[130] Lei Tian, Ziji Liu, Li-Hao Yeh, Michael Chen, Jingshan Zhong, and Laura Waller. Computational illumination for high-speed in vitro fourier ptychographic microscopy. *Optica*, 2(10):904–911, Oct 2015.

[131] Lei Tian and Laura Waller. 3d intensity and phase imaging from light field measurements in an led array microscope. *Optica*, 2(2):104–111, Feb 2015.

[132] Lei Tian and Laura Waller. Quantitative differential phase contrast imaging in an led array microscope. *Opt. Express*, 23(9):11394–11403, May 2015.

[133] Lei Tian, Jingyan Wang, and Laura Waller. 3d differential phase-contrast microscopy with computational illumination using an led array. *Opt. Lett.*, 39(5):1326–1329, Mar 2014.

[134] Peter J Verveer, Jim Swoger, Francesco Pampaloni, Klaus Greger, Marco Marcello, and Ernst H K Stelzer. High-resolution three-dimensional imaging of large specimens with light sheet–based microscopy. *Nature Methods*, 4(4):311–313, March 2007.

[135] Felipe Vico, Leslie Greengard, and Miguel Ferrando. Fast convolution with free-space green's functions. *Journal of Computational Physics*, 323:191 – 203, 2016.

[136] Curtis R. Vogel and Mary E. Oman. Iterative methods for total variation denoising. *SIAM J. Sci. Comput.*, 17:227–238, 1996.

[137] Olga V Volodina and https://pnojournal.wordpress.com/2022/07/01/volodina-3/. Formation of future teachers' worldview culture by means of foreign-language education. *P Sci Edu*, 57(3):126–159, July 2022.

[138] Zhuo Wang, Larry Millet, Mustafa Mir, Huafeng Ding, Sakulsuk Unarunotai, John Rogers, Martha U. Gillette, and Gabriel Popescu. Spatial light interference microscopy (slim). *Opt. Express*, 19(2):1016–1026, Jan 2011.

[139] Emil Wolf. Three-dimensional structure determination of semi-transparent objects from holographic data. *Optics Communications*, 1(4):153–156, 1969.

[140] Zihui Wu, Yu Sun, Alex Matlock, Jiaming Liu, Lei Tian, and Ulugbek S. Kamilov. SIMBA: Scalable inversion in optical tomography using deep denoising priors. *IEEE Journal of Selected Topics in Signal Processing*, 14(6):1163–1175, oct 2020.

[141] Guoan Zheng, Roarke Horstmeyer, and Changhuei Yang. Wide-field, high-resolution fourier ptychographic microscopy. *Nature Photonics*, 7(9):739–745, jul 2013.

[142] Guoan Zheng, Christopher Kolner, and Changhuei Yang. Microscopy refocusing and dark-field imaging by using a simple led array. *Opt. Lett.*, 36(20):3987–3989, Oct 2011.

[143] Kevin C. Zhou and Roarke Horstmeyer. Diffraction tomography with a deep image prior. *Optics Express*, 28(9):12872, April 2020.

[144] Kevin C. Zhou and Roarke Horstmeyer. Diffraction tomography with a deep image prior. *Opt. Express*, 28(9):12872–12896, Apr 2020.

[145] Pan Zhou and Jiashi Feng. Understanding generalization and optimization performance of deep cnns. *CoRR*, abs/1805.10767, 2018.

[146] Warren R Zipfel, Rebecca M Williams, and Watt W Webb. Nonlinear magic: multiphoton microscopy in the biosciences. *Nature Biotechnology*, 21(11):1369–1377, oct 2003.