



**HAL**  
open science

# Time-resolved multispectral optical tomography for reconstruction of depth- resolved changes in oxy- and deoxyhemoglobin

David Orive Miguel

► **To cite this version:**

David Orive Miguel. Time-resolved multispectral optical tomography for reconstruction of depth-resolved changes in oxy- and deoxyhemoglobin. Signal and Image processing. Université Grenoble Alpes, 2019. English. NNT : 2019GREAT108 . tel-02411550v2

**HAL Id: tel-02411550**

**<https://theses.hal.science/tel-02411550v2>**

Submitted on 9 Jul 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

**DOCTEUR DE LA**

**COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES**

Spécialité : SIGNAL IMAGE PAROLE TELECOMS

Arrêté ministériel : 25 mai 2016

Présentée par

**David ORIVE MIGUEL**

Thèse dirigée par **Jérôme MARS**, Professeur, G-INP  
et codirigée par **Lionel HERVÉ**, Ingénieur-Chercheur, CEA Grenoble et  
**Laurent CONDAT**, Chargé de recherche, CNRS

préparée au sein du **Laboratoire CEA/LETI** dans l'**École Doctorale**  
**Electronique, Electrotechnique, Automatique, Traitement du Signal**  
**(EEATS)**

**Tomographie optique multispectrale résolue  
en temps pour la reconstruction de  
changements en profondeur de l'oxy- et de la  
déoxyhémoglobine**

**Time-resolved multispectral optical  
tomography for reconstruction of depth-  
resolved changes in oxy- and  
deoxyhemoglobin**

Thèse soutenue publiquement le 8 octobre 2019,  
devant le jury composé de :

**Monsieur Alessandro TORRICELLI**

Professeur, Dept. of Physics, Politecnico di Milano, Président

**Madame Anabela DA SILVA**

Directrice de Recherche, Institut Fresnel Marseille, CNRS, Rapportrice

**Monsieur Bruno MONTCEL**

Maître de Conférences - HDR, Université Claude Bernard Lyon 1,  
Rapporteur

**Monsieur Antonio PIFFERI**

Professeur, Dept. of Physics, Politecnico di Milano, Examineur, abs.

**Monsieur Lionel HERVE**

Ingénieur-Chercheur, CEA Grenoble, Co-directeur de thèse, Examineur

**Monsieur Laurent CONDAT**

Chargé de Recherche, GIPSA-Lab, CNRS, Co-directeur de thèse, Invité

**Monsieur Jérôme MARS**

Professeur, Grenoble-INP, GIPSA-Lab, Directeur de thèse, Examineur







COMMUNAUTÉ UNIVERSITÉ GRENOBLES ALPES

---

**Time-resolved multispectral optical  
tomography for reconstruction of  
depth-resolved changes in oxy- and  
deoxyhemoglobin**

---

David ORIVE-MIGUEL

*Thesis director:* Prof. Jérôme MARS

*Thesis co-director:* Dr. Lionel HERVÉ

*Thesis co-director:* Dr. Laurent CONDAT

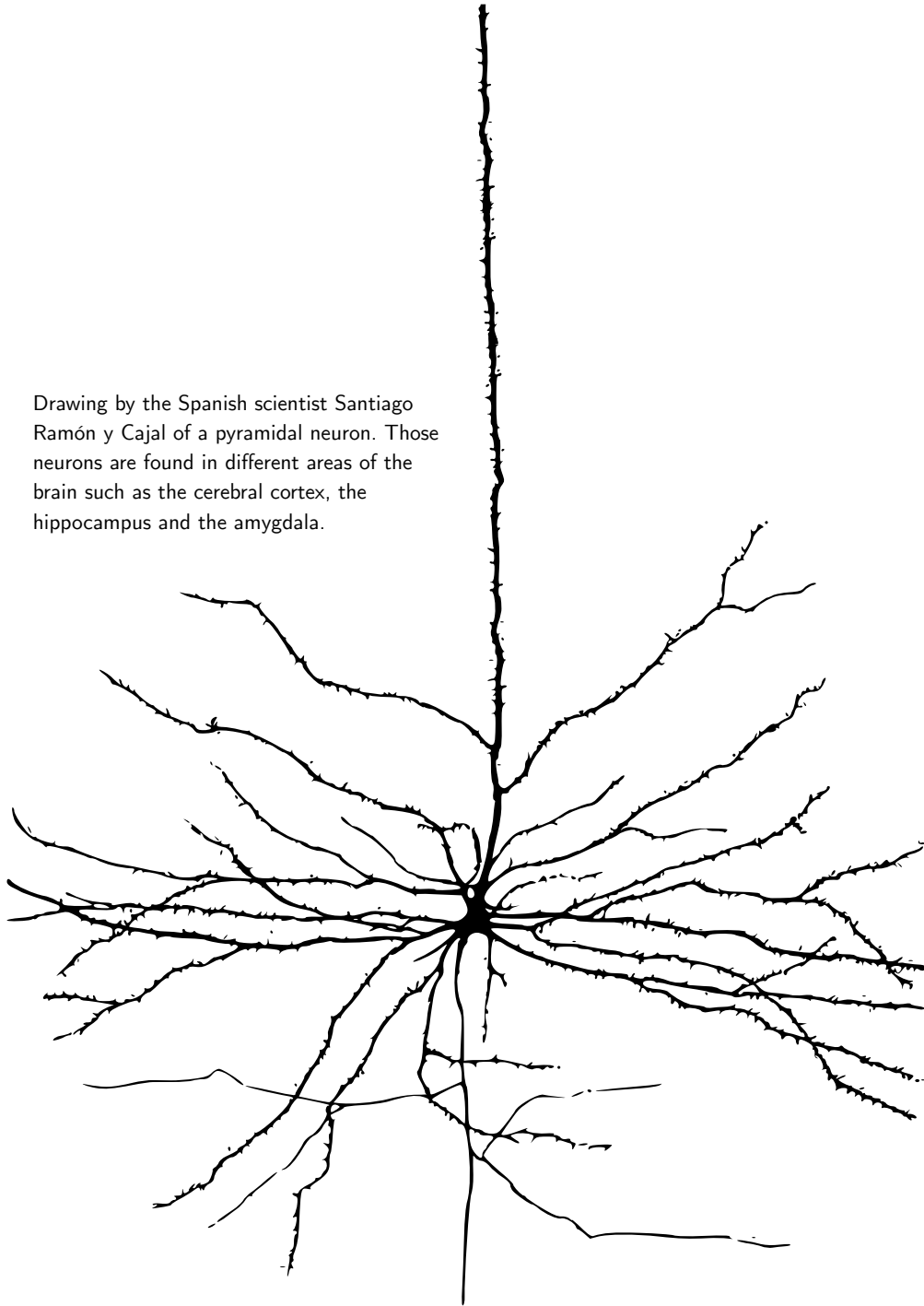




«El estudiante se encuentra con la ciencia ya hecha, como con una serranía que se levanta ante él y le cierra su camino vital. En el mejor caso, la serranía de la ciencia le gusta, le atrae, le parece bonita, le promete triunfos en la vida. Pero nada de esto tiene que ver con la necesidad auténtica que lleva a crear la ciencia. La prueba de ello está en que ese deseo general de saber es incapaz de concretarse por sí mismo en el deseo estricto de un saber determinado.»

José Ortega y Gasset, *Unas lecciones de Metafísica*.

Drawing by the Spanish scientist Santiago Ramón y Cajal of a pyramidal neuron. Those neurons are found in different areas of the brain such as the cerebral cortex, the hippocampus and the amygdala.





## Acknowledgements

First things first, I would like to acknowledge the European Union for funding this thesis. I really appreciate the effort that is being done by the European Union member countries to make European research stronger. Namely, I would like to note that this thesis has received funding from the European Union's Horizon 2020 Marie Skłodowska–Curie Innovative Training Networks (ITN–ETN) programme, under grant agreement no. 675332 BitMap.

BitMap ITN programme was an invaluable opportunity for me. Under this project, I could do research, receive high–level training and assist to many conferences around the world. Particularly, I would like to thank to Prof. Hamid Dehghani from University of Birmingham for accepting me at his lab during two intense weeks. I would also like to acknowledge Dr. Heidrun Wabnitz for admitting me at Physikalisch–Technische Bundesanstalt in Berlin for two weeks and having many interesting discussions about so many aspects of diffuse optics. Thanks also to PicoQuant company and its members for giving me training in company management and intellectual property topics. Moreover, I would like to thank to all the people from the Department of Biomedical Optics at Politecnico di Milano; in particular, to Dr. Laura Di Sieno and Dr. Alberto Dalla Mora for helping me so much during the experiments (*Grazie mille ragazzi!*). I do not forget the main actors from BitMap: the PhD students! It was very rewarding to collaborate and discuss with you about science. But even better was to share so many experiences (and beers ;) and have so much fun with you. Thank you for everything!

En ce qui concerne mon laboratoire au CEA–Leti, je voudrais remercier le soutien que j'ai reçu depuis le début de ma thèse. J'exprime également ma gratitude à tous les membres du laboratoire et aux personnes que j'ai eu l'occasion de rencontrer au cours de ces trois années. Remarque spéciale pour ceux du domaine de l'optique diffuse avec qui j'ai partagé et discuté de nombreuses idées de recherche. Je voudrais également remercier les chefs de laboratoire que j'ai eu pendant les périodes LISA et LS2P.

En dehors du travail, j'ai rencontré tellement de gens au cours de ces trois années qu'il serait impossible de les remercier tous individuellement. Pour ceux qui m'ont fait profiter et m'amuser, merci beaucoup pour tout. Remarque spéciale aux membres du club de kick–boxing BKB38 pour trois années de dur labeur et de sparring ensemble. Merci également à Echiquier Grenoblois de m'avoir accueilli dans son équipe et d'avoir participé à de nombreux tournois dans la région.

About my personal Ivory tower, many thanks to the writers that enlightened me, made me to reflect more critically and to watch the world from a different perspective; namely, thanks to A. Escotado, M. Unamuno, S. Ramón y Cajal, and many thinkers from the classical Greco–Roman period.

Back to earth, during these three years, I had the pleasure to be guided by three wise people. Laurent Condat gave me many technical advices and career development suggestions from the beginning and Jérôme Mars guided me during the whole thesis and gave me a lot of support when needed. Without your help this work could not have been possible, thank you so much for your support! Moreover, Lionel Hervé was my mentor at CEA during this journey. I cannot express how many things I learnt from you and how many interesting conversation we had. I am still fascinated by how do you manage to have original and

different ideas in so many topics. Thank you to all of you for giving me so much freedom during my PhD!

Por último, pero no por ello menos importante, me gustaría agradecer a mi familia todo el apoyo recibido durante este largo periodo. Especialmente, a mis padres y a mi hermano por preocuparos de mí. Aunque a veces me cueste expresarlo, sin vuestro esfuerzo, lucha y cariño todo esto no hubiese sido posible. Muchísimas gracias :)

---

*Gaizki esanak barkatu eta ondo esanak gogoan hartu.*

# Contents

|   |           |
|---|-----------|
| <b>1. Introduction</b>  | <b>1</b>  |
| <b>2. Background</b>  | <b>5</b>  |
| 2.1. Optical technology: an invaluable tool for physicians . . . . .        | 5         |
| 2.2. Optics in diffusive media . . . . .                                    | 9         |
| 2.2.1. Light physics in biological tissue . . . . .                         | 9         |
| 2.2.2. Photon propagation models . . . . .                                  | 14        |
| 2.3. Diffuse optics systems and instrumentation . . . . .                   | 16        |
| 2.3.1. Light sources . . . . .  | 17        |
| 2.3.2. Optical detectors . . . . .  | 18        |
| 2.3.3. Noise . . . . .  | 19        |
| 2.3.4. Time-resolved systems revisited . . . . .                            | 20        |
| 2.4. Numerical methods for photon propagation models . . . . .              | 22        |
| 2.4.1. Monte-Carlo method . . . . .   | 22        |
| 2.4.2. Finite Element method (FEM) . . . . .                                | 24        |
| 2.4.3. Computational and implementation details of FEM . . . . .            | 28        |
| 2.5. Diffuse optical tomography . . . . .                                   | 32        |
| 2.5.1. Basics of optical properties reconstruction . . . . .                | 34        |
| 2.5.2. Reconstruction using continuous-wave systems . . . . .               | 38        |
| 2.5.3. Reconstruction using frequency-resolved systems . . . . .            | 39        |
| 2.5.4. Reconstruction using time-resolved systems . . . . .                 | 39        |
| 2.5.5. Measurements into the reconstruction algorithms . . . . .            | 42        |
| 2.6. Clinical perspectives: applications to neuromonitoring . . . . .       | 44        |
| 2.6.1. Traumatic brain injuries . . . . .                                   | 44        |
| 2.6.2. Newborn hypoxia . . . . .  | 45        |
| 2.6.3. Functional imaging . . . . .   | 45        |
| <b>3. Study of datatypes for time-resolved diffuse optical tomography</b>   | <b>47</b> |
| 3.1. Introduction . . . . .   | 47        |
| 3.2. Theoretical development and analysis of new temporal windows . . . . . | 48        |
| 3.2.1. A short review of datatype-based reconstruction . . . . .            | 49        |
| 3.2.2. A novel method to compute temporal windows . . . . .                 | 50        |
| 3.2.3. Temporal windows analysis . . . . .                                  | 52        |
| 3.2.4. Noise influence . . . . .  | 57        |
| 3.3. Methods . . . . .  | 59        |



|           |  |            |
|-----------|--|------------|
| 3.3.1.    | Numerical simulations . . . . .  | 59         |
| 3.3.2.    | Quantitative evaluation metrics . . . . .  | 61         |
| 3.4.      | Results and discussion . . . . .   | 62         |
| 3.4.1.    | Comparing state-of-the-art windows with Tukey and Gaussian windows . . . . .     | 62         |
| 3.4.2.    | Comparing windows and frequency-based reconstruction . . . . .                   | 67         |
| 3.5.      | Conclusions . . . . .  | 68         |
| <b>4.</b> | <b>Total variation regularization for diffuse optical tomography</b>             | <b>71</b>  |
| 4.1.      | Regularization for DOT inverse problem . . . . .                                 | 71         |
| 4.2.      | The gradient operator in unstructured meshes: two different approaches . . . . . | 72         |
| 4.2.1.    | Finite element representation . . . . .  | 72         |
| 4.2.2.    | Graph-based representation . . . . .   | 75         |
| 4.3.      | Minimization of TV-regularized DOT inverse problems . . . . .                    | 76         |
| 4.3.1.    | The optimization problem . . . . .   | 76         |
| 4.3.2.    | Majorize-minimize algorithm . . . . .  | 77         |
| 4.4.      | Experiments . . . . .  | 81         |
| 4.4.1.    | Quantitative evaluation metrics . . . . .  | 81         |
| 4.4.2.    | Two-dimensional simulations . . . . .  | 83         |
| 4.4.3.    | Three dimensional head simulations . . . . .                                     | 85         |
| 4.4.4.    | Solid phantoms . . . . .   | 89         |
| 4.5.      | Conclusions . . . . .  | 91         |
| <b>5.</b> | <b>In-vivo experiments with a time-resolved diffuse optical system</b>           | <b>93</b>  |
| 5.1.      | Experimental setup . . . . .   | 93         |
| 5.1.1.    | Probe geometry analysis . . . . .  | 94         |
| 5.2.      | Methods . . . . .  | 99         |
| 5.2.1.    | Data preprocessing . . . . .   | 99         |
| 5.2.2.    | Data analysis . . . . .  | 100        |
| 5.2.3.    | Tomographic reconstructions . . . . .  | 101        |
| 5.3.      | Arm occlusion experiments . . . . .  | 104        |
| 5.3.1.    | Venous occlusions . . . . .  | 104        |
| 5.3.2.    | Arterial occlusions . . . . .  | 107        |
| 5.4.      | Motor cortex activation experiments . . . . .                                    | 121        |
| 5.4.1.    | Experiment description . . . . .   | 121        |
| 5.4.2.    | Results and discussion . . . . .   | 124        |
| 5.5.      | Conclusions . . . . .  | 139        |
| <b>6.</b> | <b>The BitMap open dataset</b>   | <b>141</b> |
| 6.1.      | Open Data management and FAIR principles . . . . .                               | 142        |
| 6.2.      | Standardization . . . . .  | 143        |
| 6.3.      | The diffuse optics systems campaign . . . . .                                    | 144        |
| 6.3.1.    | The European diffuse optics standards . . . . .                                  | 144        |
| 6.3.2.    | The instruments involved in the campaign . . . . .                               | 145        |
| 6.4.      | Open Data action organization . . . . .  | 146        |
| 6.4.1.    | The coordination of institutions . . . . .                                       | 146        |
| 6.4.2.    | Shared Near Infrared File (SNIRF) Format Specification . . . . .                 | 147        |
| 6.4.3.    | The open data repository . . . . .   | 150        |
| 6.5.      | Conclusions and ongoing work . . . . .   | 150        |

|  |            |
|--|------------|
| <b>7. Final conclusions and perspectives</b>   | <b>151</b> |
| <b>Appendix A. Publications</b>  | <b>153</b> |
| A.1. Published work . . . . .  | 153        |
| A.2. Papers in the pipeline . . . . .  | 154        |
| <b>Appendix B. Implementation details for photon transport Monte-Carlo method</b>    | <b>155</b> |
| B.1. Pseudo-random number generation following Henyey-Greenstein distribution        | 155        |
| B.2. Photon travelling step size . . . . .   | 156        |
| <b>Appendix C. Brief notes on vector calculus</b>                                    | <b>157</b> |
| <b>Appendix D. Generalized least squares and datatypes</b>                           | <b>159</b> |
| D.1. From ordinary generalized least squares . . . . .                               | 159        |
| <b>Appendix E. Diffusion equation and Born approximation revisited</b>               | <b>161</b> |
| E.1. Born approximation for diffusion coefficient heterogeneities . . . . .          | 161        |
| E.2. Dirac delta shift effect . . . . .  | 162        |
| <b>Appendix F. Dispersion calculations for several temporal windows</b>              | <b>163</b> |
| F.1. Gaussian window . . . . .   | 163        |
| F.2. Laplacian distribution . . . . .  | 164        |
| F.3. Rectangle function . . . . .  | 165        |
| F.4. Mellin-Laplace moments . . . . .  | 166        |
| <b>Appendix G. Fourier transform influence on noise</b>                              | <b>169</b> |
| <b>Appendix H. Extra time series results for motor cortex activation experiments</b> | <b>171</b> |
| <b>Bibliography</b>  | <b>203</b> |
| <b>Glossary</b>  | <b>205</b> |
| <b>Résumé en français</b>  | <b>207</b> |



# Chapter 1

## Introduction

**T**HIS doctoral thesis manuscript is a summary of my activity as a researcher apprentice over three years at CEA–Leti laboratories. In this part, I briefly introduce the reader to the content of diffuse optics field, its significance and most promising applications. After, I put my thesis in the framework of BitMap European project. I also describe the state-of-the-art of the community and CEA–Leti labs before my arrival and the open problems/questions that they were facing at that time. Finally, I introduce the content in each of the chapters, I explain the logical connection between them and I describe how the research presented in those chapters try to shed some light on the open questions I had to face at the beginning of my research journey.

The European Commission (EC) has selected photonics and its biomedical applications as one of the Key Enabling Technologies to improve the competitiveness of European Union in the markets [168]. In this context, big efforts are being done by European labs working at different biophotonic fields. Diffuse optics spectroscopy and imaging is one of those promising young research fields. Although, first works were published in the seventies, the foundations of the field were built in the nineties and the beginning of this century. Since those days, a lot of progress have been made and diffuse optics technology has improved considerably both in terms of hardware and data analysis. Regarding the former, now there are faster pulsed sources, larger area detectors with improved accuracy or wearable devices, just to mention a few. About the latter, quantification of optical properties have been improved, apriori information has been incorporated, sensitivity to superficial layers have been corrected and algorithms are much faster, among other improvements. Moreover, clinical applications have been widely investigated, where continuous and non-invasive neuromonitoring of adult and newborn patients is one of the most promising.

BitMap (Brain injury and trauma monitoring using advanced photonics) is an Initial Training Network supported by the EC [137] in accordance with the motivations explained above. The network is composed of 15 Early Stage Researchers, funded with Marie Skłodowska–Curie fellowships, working at leading European academic, industrial and clinical institutions. The core idea of the network is to gather PhD students with diverse backgrounds and to collaborate together to address the key technological and clinical challenges in photonic–based neurocritical care. The students have been working in different aspects of diffuse optics for neuromonitoring such as hardware development, data analysis or clinical validation. Myself, I had the opportunity to be part of this network and to contribute to improve data

analysis and tomographic reconstruction in diffuse optics. As I will show in the next chapters, during part of my research work, I collaborated with several partners from the network.

I did most of my work at CEA–Leti laboratories in Grenoble (France). Before my arrival, CEA–Leti had already done an outstanding work in the field of tomographic reconstruction using diffuse optical technology. They have an important presence in the field and are well-known because of their huge efforts in the development of time-resolved and fluorescence tomography [101, 181, 100, 26] and important advancements in small animal imaging [120, 119, 49]. They also did some preclinical tests on macaque brains for white matter lesion detection [172]. Nevertheless, they did not have experience in the application of diffuse optical tomography technology for human brain imaging. Contrary to what one may think, I found this situation very stimulating because I had a lot of freedom to do research, try new ideas and apply some of the techniques already developed at CEA–Leti into a new domain. During my PhD research, I had to face the following challenging open questions and problems:

1. Can the state-of-the-art of diffuse optical tomography algorithms be improved in terms of absorption quantification and resolution in depth? Although Mellin–Laplace moments are widely used in time-resolved diffuse optical tomography, their main weaknesses are the correlation and low temporal resolution that they have at high orders. Are there other datatypes that can be computed fast and diminish those drawbacks?
2. As it is known in the functional near-infrared spectroscopy community, one of the most common problem in brain imaging is the influence that scalp layer has on the measurements. Since this layer is very close to the sources and detectors it has a big influence in the measurements. The questions that arise are: how can be compensated the larger sensitivity of shallow layers? What are the different approaches to normalize the sensitivity in depth? How can these methods be applied for brain tomographic imaging without losing important information about brain activation?
3. The developed instruments and algorithms need to be adapted for brain imaging. But, what probe geometries are better for brain imaging? How can the algorithms be adapted to reconstruct the hemoglobin oxygenation and deoxygenation in the brain at deep layers? What assumption can be done and what type of regularization methods can be applied in this problem? Moreover, although brain conditions are not constant over time due to the diverse physiological processes that take place, can brain resting periods be used to regularize the problem in time?
4. The last topic of my thesis concerns about the translation of academic research into the clinics and industry. In the European Union, huge effort is being done to push those technologies into the market through promotion of standardization. Therefore, some of the questions that arise are: can standardization of instruments assessment drive biophotonics technologies into the clinics? How can standardization be promoted in the field? Could the publishing of Open Data promote standardization? How Open Data needs to be published to promote its reusability in the long-term?

In this manuscript, I try to answer these questions and to open new paths for future research. The manuscript is organized in six chapters, including this introduction. In the second chapter, I introduce to the reader into the biomedical photonics world. First, I give an introduction to biomedical imaging and the role that biomedical optics has on it. After, I describe the physics of light in diffusive media and the different photon propagation models that are used in the literature. Then, the mathematics behind numerical methods for photon

propagation models and diffuse optical tomography are explained. I also discuss about the different types of instrumentation that are found in diffuse optics labs. At the end of the chapter, I describe three major applications of diffuse optics into the clinics.

In the third chapter, I report the analysis I did about different kind of datatypes for time-resolved diffuse optical tomography. First, I propose a novel and efficient technique for computing a larger set of temporal windows. Then, I analyze the Gaussian and Tukey windows by computing their dispersion levels at time and frequency domains. These new temporal windows are compared against the well-known Mellin–Laplace moments and Fourier transform datatype. Moreover, I perform a theoretical analysis of the noise correlation that arises from these datatypes. At the end, some numerical results are presented for spherical inclusions at different depth levels.

In the fourth chapter, I describe the work I did in collaboration with PhD student Wenqi Lu and her supervisor Dr. Iain Styles from the University of Birmingham. In this work, we study the effect of total-variation regularization in diffuse optical tomography. We proposed two different approaches to impose total-variation solutions into diffuse optical tomography problems. One of them is based on finite element approach and the other in a graph based method. First, both methods mathematical background is described in detail. Then, the regularization methods are validated in numerical simulations using a head model. Finally, the methods are tested against experimental measurements from a cylindrical phantom using a continuous-wave system.

In the fifth chapter, I apply the techniques and algorithms developed at previous chapters into in-vivo experiments. First, I describe the dual-wavelength (670 nm and 820 nm) optical instrument composed by two probes containing three fibers and four silicon photomultiplier detectors each. Then, I show the tomographic simulations I performed to analyse the best geometry of source-detector positions in terms of tomographic localization and resolution. After, I detail the data analysis and tomographic reconstruction techniques for in-vivo imaging. Later, I describe arm occlusion experiments that I performed to validate the proposed optical system. Finally, I show the results for brain motor cortex action experiments with three human subjects. The results show that it is possible to monitor with one-second resolution the motor cortex activation and that spatial and in depth information can also be retrieved.

In the sixth chapter, I detail the effort that BitMap network is doing to promote standardization of diffuse optics and promote Open Data philosophy in the field. I describe the work that was done to measure up to 28 instruments from 8 different European institutions using three well-agreed international protocols. After, I detail my ongoing work on gathering all the measurements from different institutions and building an Open Data dataset with those measurements. I describe which format I am using to organize all the data and I highlight the importance of making this dataset findable, accesible, interoperable and reusable in the long-term.

After the main chapters, I included an appendix where I list the work I published both in journals and conferences during these three years. Moreover, I also mention the papers that I have in the pipeline. After, a few more appendices are included that support the work I presented in the previous chapters.



## Background

**I**N this chapter, I will describe the mathematical and physical foundations that were used during my Ph.D. research. Emphasis is put on the numerical modelling of photon propagation in diffusive medium. An analysis of already developed reconstruction techniques is also given. The techniques described here were also put in their scientific and clinical context in order to capture the whole picture and potential applications of this thesis work.

### 2.1. Optical technology: an invaluable tool for physicians

When a retrospective view of last century medicine is taken, undoubtedly many of the greatest contributions are related to imaging. Before the discover of  $X$ -radiation in 1895 by the German physicist Wilhelm Röntgen [188] physicians could only make a non-invasive prognosis of broken bones, musculoskeletal pain or tumor detection based on palpation. Therefore, since most manual techniques depend highly on the experience of the physicians, that made the accuracy of such prognosis very uncertain. The other possibility was to go inside the patient to observe directly the source of the illness but in most cases this was risky and patient recovery took longer.

After the appearing of  $X$ -rays, the context of medicine changed dramatically; bone structure and some tissues could be seen in great detail without the need of making any incision. Sixty years after, ultrasound based medical imaging systems [115] were developed to image soft tissues like organs, muscles and blood vessels. Around the seventies, mathematical methods and advances in computer sciences were used to create new imaging techniques. A great example is computed tomography (CT), first developed by Sir Godfrey Hounsfield [104], that reconstructs patient 3D images using  $X$ -rays projections taken at different angles in combination with the mathematical theory of the Radon transform. Magnetic resonance imaging (MRI) is another example of a medical imaging technique that relies on mathematics and computing to recover 3D images. Approximately at the same time as computing was introduced to medical imaging, new techniques based on injecting external chemical agents were also developed; the idea is to attach that agents into tumors or organs in order to enhance their image quality. Examples of this technology are single-photon emission computed tomography (SPECT), positron emission tomography (PET) and fluorescence imaging that



uses radioisotopes, positron emitting isotope and fluorescent stains, respectively, in order to build a 3D image of human body parts.

Nowadays, medical imaging is in constant development, not only by the improvement of current imaging devices, but also by including new technologies that will help physicians to make more accurate decisions. In the former case, improvements are done in terms of increasing the spatial and temporal resolution of already existing techniques [224, 191, 232, 246, 176]. Moreover, multi-modal imaging is also supported, that is, to include several already existing imaging tools into one device so the strengths of one technology can overcome the weakness of the other one and vice versa; for example, ultrasound technology have been included into diffuse optical tomography systems to predict the outcome of chemotherapy into breast cancer patients [255, 242] or to assess thyroid tissue [132]. Huge effort is also being done to accelerate the processing of images by including high performance GPUs [204, 66] and improving the reconstruction algorithms [158, 41]. Regarding the development of novel technologies, several quantitative biomarkers associated with various diseases are being investigated [180, 2]. Great advances are also being made by applying artificial intelligence to, for example, perform automatic segmentation of organs or to support clinicians in the classification of benign and malignant tumors [86, 236]. In addition, since current medical systems allow to record a large amount of medical data from a patient, Big Data approaches can also be included into the artificial intelligence algorithms [153, 252]. Finally, important improvements are being done in the development of image-guided surgery systems. Those systems do not only make easier and more precise the work of surgeons, but also permit to do small incisions, thanks to minimal invasive surgeries (also known as *keyhole* surgeries), which consist on an endoscopy tube with a video camera and small surgical tools, that way the risks of infection or times of recovery are decreased significantly [212, 184].

On a daily basis of hospitals and medical centers, a large range of medical imaging technologies are used: from a simple *X*-ray radiography to assess how badly a bone is broken, to detect breast cancer thanks to mammographies or to investigate how different parts of the brain work, by using functional MRI (fMRI). These technologies have different properties and characteristics which make them suitable depending on the clinical situation. The main characteristics that a medical imaging technology can have are the following:

- **Spatial resolution:** It is defined as the minimum distance where two objects close to each other can still be recognized as separated. Some typical values for different imaging techniques are shown at Table 2.1. Note that these values are dependent on several factors like individual characteristics of the system, artifacts or post-processing algorithms that were used.
- **Contrast resolution:** It is the ability to distinguish between intensity differences of an image. Contrast is important to discriminate between different tissues. Performance in terms of contrast resolution for several medical imaging systems are given at Table 2.1.
- **Temporal resolution:** This property determines how fast a system is performing an acquisition. For technologies applied to real-time situations, this is a critical aspect since the temporal resolution should be as high as possible. A paradigmatic case is fMRI where different areas of brain activation should be recorded in delays of a few seconds in order to determine how the brain behaves when a patient performs a given task. Approximated values for several systems are given at Table 2.1.
- **Cost:** Unfortunately, the price of medical technologies is an important point when the availability of a technology is assessed. Some technologies like MRI, CT or SPECT are

expensive and their availability is not only compromised in third-world countries but also in some local hospitals of developed countries. For that reason, alternative imaging technologies have been developed, such as ultrasonography or optical tomography, which have a lower cost and are largely available to the population.

- **Counter-indications:** Although the frequency of use is limited by the cost of the technology it can also be restricted by intrinsic or patient-dependant counter-indications. Classic examples are *X*-rays or SPECT technologies which are not recommended to be applied to a given patient on a constant basis due to health issues that large amounts of radiation can lead to. Moreover, for some patients like newborn babies special restrictions are applied such as motion artifacts or radiation tolerance.
- **Complexity:** Some technologies such as CT or MRI are so sophisticated they require highly trained technicians not only for the maintenance but also for results analysis and use. This complexity can be tolerated for some technologies since the intervention of the physicians is not required, e.g. CT or MRI. Nevertheless, this does not hold for all techniques; for example, the technology used for real-time fluorescence image-guided surgeries has to be controlled by the surgeons during the intervention and, therefore, it must be easy to use.

| Technology            | Spatial resol.       | Contrast resol. | Temporal resol.   |
|-----------------------|----------------------|-----------------|-------------------|
| CT                    | 0.1-0.6 mm           | Low to moderate | 100 ms            |
| MRI                   | 0.1-2 mm             | High            | 50 ms             |
| PET                   | 4-10 mm              | Very high       | 5 min             |
| SPECT                 | 4-15 mm              | Very high       | 15 min            |
| Ultrasonography       | 2-3 mm               | Low to moderate | 40 ms             |
| Photoacoustic imaging | 20-100 $\mu\text{m}$ | Very high       | 100 $\mu\text{s}$ |
| DOT                   | 0.5-3 cm             | Very high       | 100 $\mu\text{s}$ |

Table 2.1: Comparison of most used medical imaging techniques based on spatial, contrast and temporal resolution. Data was obtained from [131, 231] and from personal experience.

Medical imaging technologies are based on different physical phenomena that are well understood and, therefore, they can also being exploited for the benefit of the society. For example, MRI technology is based on the understanding of magnetic fields, radio waves, and how atomic nuclei behave under strong magnetic gradients. The knowledge of these fields makes it possible for the society to exploit and use them for its own benefit. Light is one of those physical phenomena that has been studied since the Ancient Greece and large advances have been made from the XVII century to nowadays. The knowledge of classical optics allowed to create the first glasses to correct several conditions of the eye. Since then, the theory behind optics has been applied to support physicians in different ways. One of the first medical tools with optical technology was the otoscope, which is used to visualize the ear canal and eardrum and was started to be commercialized around the nineteenth century. After the otoscope, the ophthalmoscope and the retinoscope were invented in order to examine the interior of the eyeball and the retina respectively. Around the sixties, optical fibers were developed to carry light and images from places that were far away. Optical fibers were rapidly applied to medicine, since they are small and flexible, so they can be easily fitted into the human body. Glass optical fibers can be assembled together; these set of optical fibers can be incoherent or coherent. In the first case, their position at the

beginning of the bundle and at the end is random (usually they do not coincide) therefore they are used for illumination or transmission of signals. On the coherent case, each fiber has the same relative position at the beginning and the end of the bundle so this type of optical fiber bundles can be used for real-time image transmission. Nowadays, endoscopes are constructed using these fiber optics bundles and allow the clinician to visually inspect several organs and body cavities in real-time (see Figure 2.1a and Figure 2.1b for a bronchoscopy and stomach endoscopy cases respectively). Another application of optical technology into medicine is near-infrared lasers to measure blood oxygenation and cardiac frequency; these devices are commonly known as oxymeters. Lasers at different wavelengths have also been used for cancer diagnosis and treatment. They have many applications in dentistry and plastic surgery. High-power lasers can also be used as a high precision ablation and cutting tools. Finally, one of the most cutting-edge technologies are optical coherence tomography (OCT) and fluorescence guided surgery. In the first case, near-infrared light is used to obtain micrometer resolution images; it has been applied in ophthalmology, cardiology and oncology among many other medical fields. In the second case, fluorescent molecules are attached to tumors to better mark out the area of malign tissue.

As it was shown before, clinics is full of optical based medical devices, from the simple pulse oxymeter that measures oxygen saturation in the blood based on Beer-Lambert law to the fluorescence technology that labels different tissues during an oncology surgery. Optical based systems can be used both for imaging and to treat some diseases. The advantages of optics based medical systems are the following:

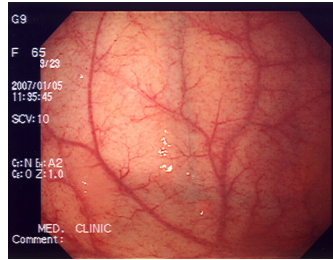
- **Scalability, portability and flexibility:** Most of the optical systems are highly scalable so its size can be adapted to the clinical constraints. Moreover, since the development of optical fibers, many of those systems are very flexible and they can be introduced safely into human body to reach several organs and human body cavities. In some cases, portable or wearable devices can also be developed.
- **Interaction with tissue:** Since photons have energy, they will interact with human tissue via different process such as absorption, scattering or fluorescence. Those physical interactions can be analyzed and processed in order to provide physicians with valuable biological information about a given disease or tissue.
- **Multi-modality:** Most of the optical systems can be used in combination with other medical imaging techniques. The main reason is to take advantage of optical imaging high sensitivity. For example, multi-modal systems with CT [241], PET [196], nuclear imaging [3], fluorescence and photoacoustic imaging [12, 133] have been developed.
- **Cost:** The cost of most optical-based technology is usually low compared with other medical imaging modalities. This property makes many optical technologies suitable also for underdeveloped countries. Moreover, many parts of the optical systems are reusable and disposable parts are usually not expensive. In fact, some companies are commercializing cheap enough totally disposable endoscopes.
- **High temporal resolution:** Optical systems are fast and image processing performance have improved considerably since the appearance of GPUs and new high computing capacities. Therefore, most of optical imaging techniques are suitable for real-time in-vivo applications due to its high temporal resolution rates.
- **Safety:** There are no major counter-indications when applying a light source, such as lasers, into the human body up to a power limit. In addition, optical systems are easy

to sterilize so most of their parts can be reused or, in some cases, their low cost makes them fully disposable.

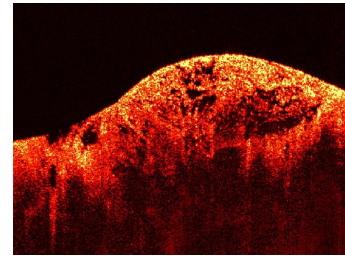
- **Training:** For most of the systems there is no need of large training in order to use the devices properly. Moreover, usually no specific assistant is needed to manipulate the instrument.



(a) Clinician doing a bronchoscopy.



(b) Image from a stomach endoscopy.



(c) Sarcoma imaged using OCT.

Figure 2.1: Examples of optical technology used in the clinics. (All photos are in public domain and taken from Wikimedia Commons).

## 2.2. Optics in diffusive media

As was explained in the previous section, many light based technologies have been already implemented successfully into the clinics. In near-infrared spectroscopy (NIRS) field pioneering work was done in 1977 by Frans F. Jöbsis [112, 113] who monitored Cytochrome c oxidase in-vivo using infrared transillumination. After that, diffuse optics field has evolved considerably and many new ideas have been developed. In the following subsections theoretical background on the physics of light in high scattering media will be given, as well as the biochemical composition of human tissue and how it affects to light will be discussed. Further on, different ways of modelling photon propagation in tissue will be addressed.

### 2.2.1. Light physics in biological tissue

It is well-known that human tissue is a highly scattering medium. In diffuse optics, scattering is defined by two parameters: (1) the frequency a photon scatters and (2) the average of the scattering angle cosine. Mathematically this can be modelled using the parameter  $\mu_s$  [ $\text{mm}^{-1}$ ] that expresses on average how many scattering events a photon suffers when it travels a distance of one millimeter (its inverse defines on average how much distance the photon travels without having a scattering event) and  $g$  [unitless], also known as the anisotropy factor, that takes a value from negative to positive one and determines the direction distribution of a photon after a scattering event. The anisotropy factor is modelled as

$$g = \int_{4\pi} \cos \theta p(\theta) d\Omega, \quad (2.1)$$

where  $p(\theta)$  [sr<sup>-1</sup>] is the normalized phase function that represents the probability per unit solid angle that a photon scatters at angle  $\theta$  relative to the incident direction. Since it is normalized, the integral over all the solid angle is one. For isotropic particles the anisotropic factor expression is simplified to

$$g = 2\pi \int_0^\pi \cos \theta p(\theta) d\theta. \quad (2.2)$$

Usually the phase function is modelled by the Henyey-Greenstein function [98] (see Figure 2.2),

$$p(\theta) = \frac{1}{4\pi} \frac{1 - g^2}{(1 + g^2 - 2g \cos \theta)^{3/2}}, \quad (2.3)$$

which has been proved valid for human tissue [23].

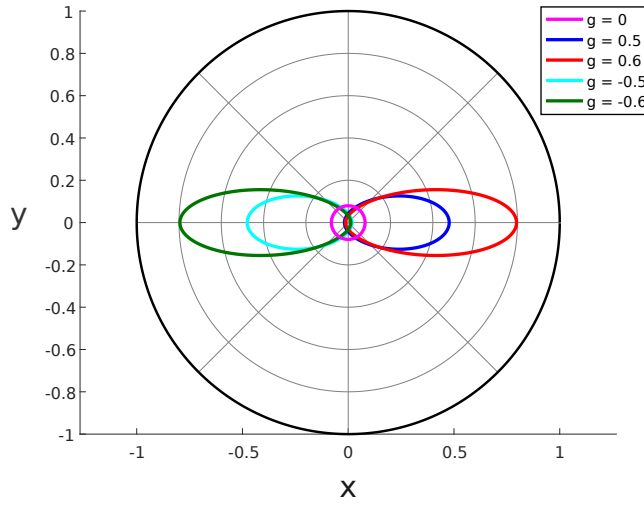


Figure 2.2: Henyey-Greenstein phase function for several anisotropic factor values. It is assumed that the photon goes from left to right though the x-axis.

The anisotropy factor of human tissue is around  $g = 0.9$ , which implies that mainly forward scattering takes place. These two parameters can be merged in one called reduced scattering parameter  $\mu'_s = \mu_s(1 - g)$ . It can be understood as an effective scattering parameter, that is, it describes diffusion of photons assuming isotropic scattering. Regarding scattering dependence on wavelength, in human tissue it has been validated experimentally that it follows a power law [154, 256, 243] and theoretical models, such as Mie theory, predict this relationship. Usually it is described as,

$$\mu'_s(\lambda) = \mu'_s(\lambda_s) \left( \frac{\lambda}{\lambda_s} \right)^{-\gamma}, \quad (2.4)$$

where  $\lambda_s = 1000$  nm and  $\gamma$  is a parameter that depends on the size of the scattering particles. A larger value of  $\gamma$  corresponds to smaller scattering particles and a lower value  $\gamma$  values to bigger particles. In [256], a summary of experimentally determined  $\gamma$  values for different human tissues can be found. At Figure 2.3, the power law dependence for several  $\gamma$  values is given.

A photon, apart from scattering, can also be absorbed when it collides with a chromophore that absorbs light at visible and near-infrared wavelengths. When a photon collides with a

chromophore, its energy is converted, usually, into thermal energy. In that case, the photon is 'lost' and the measured intensity will decrease. Absorption of a medium with just one type of compound is defined as  $\mu_a = C\varepsilon$  where  $C$  [ $M = \text{mol/L}$ ] is the molar concentration of a compound and  $\varepsilon$  [ $\text{mm}^{-1}/M$ ] is the molar extinction coefficient. If the medium is composed of  $N$  compounds then the absorption is computed by the mixture law, defined as:

$$\mu_a = \sum_{j=1}^N C_j \varepsilon_j. \quad (2.5)$$

The human body is composed of several chromophores. Hemoglobine is one of the most important chromophores efficient for the visible spectrum. It is located at the red blood cells and its main task is to transport oxygen through human body. Hemoglobine can be expressed in two states: oxygenated or deoxygenated; in the former case, hemoglobin is carrying oxygen molecules and is known as oxy-hemoglobin ( $\text{O}_2\text{Hb}$ ), in the latter case the hemoglobin is free of oxygen and is called deoxy-hemoglobin (HHb). Oxygen saturation is defined as the percentage of oxyhemoglobin compared to the total quantity of hemoglobin,  $\text{O}_2\text{Hb}/(\text{O}_2\text{Hb} + \text{HHb})$ . Detailed studies have been done to characterize the absorption of each hemoglobine state. At the end of the 20<sup>th</sup> century, Scott Prahl compiled molar extinction coefficients from several sources and did an estimation of each spectrum [178, 110] (see Figure 2.4a). The absorption of hemoglobin at shorter wavelengths is quite large. Nevertheless, at the range 600-900 nm, the absorption decreases significantly and the difference between oxy- and deoxy-hemoglobin is evident; that range is usually called the *optical window* or *therapeutic window*. At this range, light is not highly absorbed and the separation allows to deduce the concentration of each state of hemoglobin by using Beer-Lambert law [22].

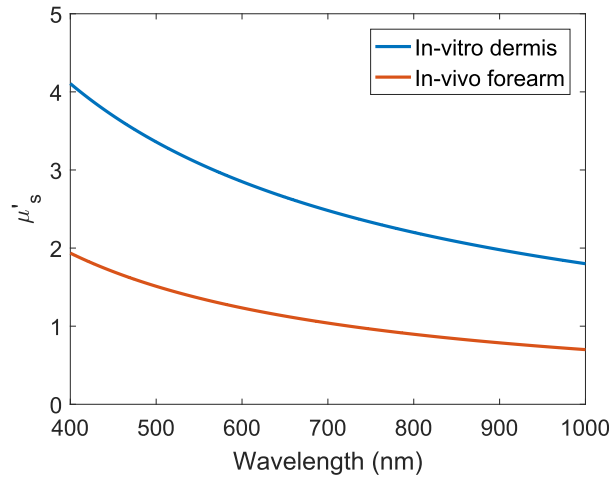
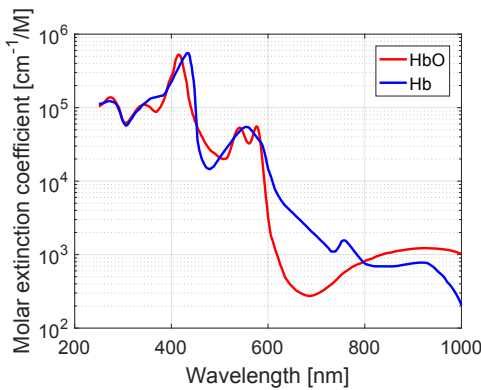


Figure 2.3: Power law dependence of scattering on the wavelength described by Equation 2.4 for two  $\gamma$  values. For the blue line in-vitro dermis properties obtained at [190, 202] were used,  $\mu'_s(1000 \text{ nm}) = 1.8$  and  $\gamma = 0.9$ . For the orange line in-vivo forearm properties from [140, 60, 221] were taken,  $\mu'_s(1000 \text{ nm}) = 0.7$  and  $\gamma = 1.11$ .

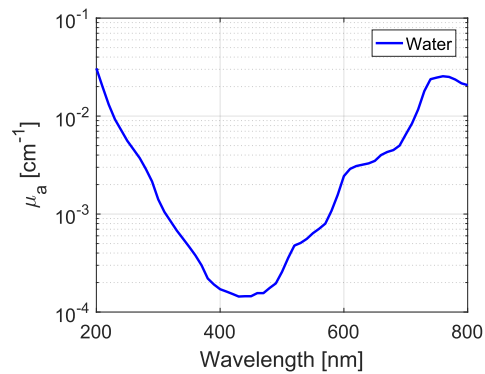
Although water is not considered a chromophore, it is still an absorber. Its absorption is quite low but since the human body has high quantities of water, the absorption can be

significant at some wavelengths. In Figure 2.4b the absorption spectrum of water is shown; at the visible spectrum (400-700 nm wavelength) the absorption is low, nevertheless at near-infrared range ( $> 700$  nm) the absorption is much stronger and it can decrease significantly the number of detected photons.

Light can penetrate large distances (in the order of meters) into the human tissue, however it is known that the probability for such an event to happen is very small and noise would hide any information that those photons could carry. The high scattering nature of human tissue constraints the average penetration depth into a few centimeters [138] in the best case. This fact imposes some limitations into the clinical applications of diffuse optics technology. However, there are still very interesting applications such as brain cortex blood oxygenation estimation [63, 146, 71], breast cancer detection [222, 223, 99, 72] or analysis of thyroid nodules [132, 122] where the needed penetration is just a few centimeters. In addition, diffuse optics technology can also be applied to small animal imaging [88, 24, 25].



(a) Molar extinction coefficient of oxy and deoxyhemoglobin at visible spectrum. Data taken from [178].



(b) Absorption of clear natural water. Data taken from [205].

Figure 2.4: Absorption spectra of hemoglobin and water at 200-800 nm wavelength.

In order to overcome the constraint that multiple scattering imposes into the penetration of light, some researchers are working on shaping the wavefront of light [247]. The basic idea is to modify the wavefront shape to suppress or control the multiple scattering of photons in order to focus light at deeper tissue layers. This is done by receiving some feedback from fluorescence or photoacoustic signals and shaping the wavefront based on a given criteria [226]. At first stages, this was achieved for static phantoms, that is, phantoms whose optical properties were constant and scattering particles did not change their position [134, 143]. Now, the wavefront shaping can be done in real-time, so light is focused on dynamic mediums where scattering particles are changing constantly [161, 53]; examples of light focusing for in-vivo tissues cases can be found at [164, 183, 165].

### Beer-Lambert law

Between the 18th and 19th century, the physicists Johann Lambert and August Beer concluded independently that if a beam of light is passed through a homogeneous and non-scattering medium then, the attenuation maintains a logarithmic relation with the absorption

of the medium. Mathematically, the decrease of light by an infinitesimal thick medium is described as:

$$dI = -I\mu_a dx, \quad (2.6)$$

where  $dx$  is the infinitesimal thickness of the layer and  $I$ , the intensity of light. After integrating through all the thickness the intensity is,

$$I = I_0 e^{-\mu_a L}, \quad (2.7)$$

where  $I_0$  is the initial intensity of the beam.

This law can also be applied to non-absorbant and low scattering mediums (single scattering is assumed,  $\mu_s L \ll 1$ ),

$$I = I_0 e^{-\mu_s L}. \quad (2.8)$$

Neither Equation 2.7 nor Equation 2.8 account for high scattering mediums since when multiple scattering occurs the path that photons travel is not the straight path  $L$  but a much larger one. For that reason, at the end of the eighties David Thomas Delpy proposed a modified version of Beer-Lambert law [57] where he introduced the mean optical path-length  $\langle L \rangle$  parameter which represents the average of the distance traveled by each photon.  $\langle L \rangle$  parameter depends on the scattering and can be determined by time-resolved optical technology. Therefore, the modified expression is,

$$I = I_0 e^{-\mu_a \langle L \rangle}. \quad (2.9)$$

Absorption measurements at different wavelengths can be used to determine the concentration of chromophores such as oxy and deoxyhemoglobin. As was expressed before, absorption due to  $N$  compounds is defined as:

$$\mu_a(\lambda_j) = \sum_{i=1}^N \varepsilon_i(\lambda_j) C_i, \quad (2.10)$$

where, obviously,  $\mu_a$  and  $\varepsilon$  depend on the wavelength  $j$  and  $C_i$  is the concentration of compound  $i$ . If several measurements are done at different wavelengths in order to make the system determined then  $C_i$  can be retrieved. By looking at Figure 2.4a, it is clear that picking wavelengths around  $\lambda_1 = 660$  nm and  $\lambda_2 = 900$  nm would be optimal selections. The concentration at 800 nm wavelength is called an isobestic point since the molar extinction coefficient is the same for oxy and deoxyhemoglobin. Although this wavelength is not useful for deriving the concentration of each hemoglobin it gives easily the total concentration and many systems include it. The system to solve is

$$\begin{bmatrix} \mu_a(\lambda_1) \\ \mu_a(\lambda_2) \end{bmatrix} = \begin{bmatrix} \varepsilon_1(\lambda_1) & \varepsilon_2(\lambda_1) \\ \varepsilon_1(\lambda_2) & \varepsilon_2(\lambda_2) \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} \quad (2.11)$$

or, equivalently,

$$- \begin{bmatrix} \frac{1}{\langle L \rangle_{\lambda_1}} \ln \left( \frac{I(\lambda_1)}{I_0(\lambda_1)} \right) \\ \frac{1}{\langle L \rangle_{\lambda_2}} \ln \left( \frac{I(\lambda_2)}{I_0(\lambda_2)} \right) \end{bmatrix} = \begin{bmatrix} \varepsilon_1(\lambda_1) & \varepsilon_2(\lambda_1) \\ \varepsilon_1(\lambda_2) & \varepsilon_2(\lambda_2) \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix}. \quad (2.12)$$

where  $\varepsilon_1$  and  $\varepsilon_2$  are obtained experimentally. Equation 2.10 assumes that the measurement of  $\mu_a$  was not influenced by  $\mu'_s$ , which usually does not hold. In such cases, the relative measurement  $\Delta\mu_a$  can be computed as:

$$\Delta\mu_a(\lambda_j) = \sum_{i=1}^N \varepsilon_i(\lambda_j) \Delta C_i, \quad (2.13)$$



so only the relative concentrations will be calculated, that is, the difference in concentration between the measurements done at starting time and after. In this case it is assumed that scattering is constant along measurements.

### 2.2.2. Photon propagation models

In this subsection, I define physical quantities of light and the most popular photon propagation models. The Ph.D. work presented here is highly rooted in those models.

#### Physical quantities

*Radiance*,  $L(\mathbf{x}, \hat{s}, t)$ , is defined as the radiant power per unit normal area per unit solid angle, [ $\text{Wm}^{-2}\text{sr}^{-1}$ ] that depends on six variables: space ( $\mathbf{x}$ , three spatial variables), orientation ( $\hat{s}$ , two variables) and time ( $t$ , one variable), see Figure 2.5 where the radiance received by a detector is shown.

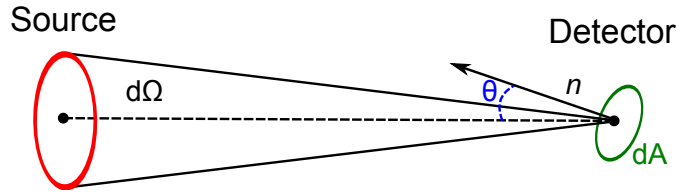


Figure 2.5: Radiant power received by a detector within a differential area  $dA$  and subtending a differential solid angle  $d\Omega$  with the source.

The *fluence rate* is defined as:

$$\Phi(\mathbf{x}, t) = \int_{4\pi} L(\mathbf{x}, \hat{s}, t) d\Omega \quad [\text{Wm}^{-2}], \quad (2.14)$$

which is the radiance integrated over the entire solid angle.

#### Radiative Transfer Equation

The Radiative Transfer Equation (RTE) is derived from the conservation of energy principle by neglecting coherence, polarization and non-linearity. The derivation presented here is based on the one described in [234].

Applying the conservation of energy inside a volume element, the following four terms must be considered. In the first place, when light beam is not collimated, the divergence of power in a unit volume element per unit time is

$$dP_1 = \frac{\partial L(\mathbf{x}, \hat{s}, t)}{\partial s} d\Omega dV = \hat{s} \cdot \nabla L(\mathbf{x}, \hat{s}, t) d\Omega dV, \quad (2.15)$$

where  $d\Omega$  and  $dV$  are the differentials of solid angle and volume respectively.

The second term is related to the power loss due to absorption of photons or scattering into another direction,

$$dP_2 = (\mu_a + \mu_s) L(\mathbf{x}, \hat{s}, t) d\Omega dV = \mu_t L(\mathbf{x}, \hat{s}, t) d\Omega dA, \quad (2.16)$$

where  $dA$  is the differential area and  $\mu_t = \mu_a + \mu_s$ .

Regarding the third term, scattering can decrease power in a given direction (modelled by the second term previously seen) but can also increase when the scattered photons from other directions that take the considered direction after a scattering event,

$$dP_3 = (N_s dV) \left[ \int_{4\pi} L(\mathbf{x}, \hat{s}', t) p(\hat{s}', \hat{s}) \sigma_s d\Omega' \right] d\Omega, \quad (2.17)$$

where  $N_s$  is the density of scatterers,  $N_s dV$  denotes the number of scatterers in a differential volume,  $\sigma_s$  is the scattering cross-section of a scatterer and  $p(\hat{s}', \hat{s})$  is the phase function denoting the probability that a photon with direction  $\hat{s}'$  will be scattered to direction  $\hat{s}$ . Of course, the function is normalized,

$$\int_{4\pi} p(\hat{s}', \hat{s}) d\Omega' = 1. \quad (2.18)$$

Moreover, since usually the phase function does not depend on the directions but on the angle between them, it can be simplified to

$$p(\hat{s}', \hat{s}) = p(\cos \theta), \quad (2.19)$$

where  $\theta$  is the angle between directions. Now, realizing that  $\mu_s = N_s \sigma_s$  it follows that

$$dP_3 = (\mu_s dV) \left[ \int_{4\pi} L(\mathbf{x}, \hat{s}', t) p(\hat{s}' \cdot \hat{s}) d\Omega' \right] d\Omega. \quad (2.20)$$

The last element that can produce power are the sources in the domain with units  $[\text{Wm}^{-3} \text{sr}^{-1}]$ ,

$$dP_4 = S(\mathbf{x}, \hat{s}, t) dV d\Omega. \quad (2.21)$$

It is also known that the change of energy in a volume unit element is given by:

$$dP_{tot} = \frac{1}{c} \frac{\partial L(\mathbf{x}, \hat{s}, t)}{\partial t} dV d\Omega, \quad (2.22)$$

where  $c$  is the speed of light in the medium. Therefore, the principle of energy conservation states that

$$\begin{aligned} \frac{1}{c} \frac{\partial L(\mathbf{x}, \hat{s}, t)}{\partial t} &= -\hat{s} \cdot \nabla L(\mathbf{x}, \hat{s}, t) - \mu_t L(\mathbf{x}, \hat{s}, t) \\ &+ \mu_s \int_{4\pi} L(\mathbf{x}, \hat{s}', t) p(\hat{s}' \cdot \hat{s}) d\Omega' + S(\mathbf{x}, \hat{s}, t). \end{aligned} \quad (2.23)$$

which is known as the *Radiative Transfer Equation*.

### Diffusion approximation

The Radiative Transfer Equation (Equation 2.23) is an integro-differential equation with six dimensions. This equation is usually solved using Monte-Carlo methods, which are very time-consuming (see Subsection 2.4.1). Consequently, a common approximation is to represent radiance using spherical harmonics and to assume that (1) reduced scattering coefficient is much greater than absorption ( $\mu'_s \gg \mu_a$ ) and (2) that detectors are sufficiently far from light

sources. After the introduction of spherical harmonics and making the previous assumptions, the obtained partial derivative equation (PDE) is given by:

$$\frac{1}{c} \frac{\partial \phi(\mathbf{x}, t)}{\partial t} - \nabla \cdot (D(\mathbf{x}) \nabla \phi(\mathbf{x}, t)) + \mu_a(\mathbf{x}) \phi(\mathbf{x}, t) = S(\mathbf{x}, t), \quad (2.24)$$

where  $\phi$  is the fluence rate,  $D = (3(\mu_a + \mu'_s))^{-1}$  is the diffusion coefficient and  $S(\vec{r}, t)$  is the source function. A detailed derivation can be found at [234]. In the literature, the time-dependant fluence rate  $\phi(t)$  is usually known as the temporal point spread function (TPSF). In this work, the signal computed by the Diffusion Approximation equation will be denoted as TPSF. Meanwhile, the arrival of photons histogram obtained from experiments and Monte-Carlo method will be written as photon time-of-flight distribution (DTOF). Note that under normalization, both curves are equivalent.

Every PDE must have initial and boundary conditions. For the former, zero fluence rate is assumed  $\phi(\mathbf{x}, 0) = 0$ , and for the latter a Robin type boundary condition is applied (see [195] for computational validation),

$$\phi(\mathbf{x}, t) - 2C_R D(\mathbf{x}) \frac{\partial \phi(\mathbf{x}, t)}{\partial n} = 0, \quad \mathbf{x} \in \partial\Omega, \quad t > 0, \quad (2.25)$$

where  $C_R = (2/(1 - R_0) - 1 + |\cos \theta_c|^3) / (1 - |\cos \theta_c|^2)$  is a parameter derived by the Fresnel laws being  $\theta_c = \arcsin(1/n_r)$  and  $R_0 = (n_r - 1)^2 / (n_r + 1)^2$  where  $n_r$  is the refraction index and  $\partial\Omega$  is the boundary of the domain.

Equation 2.24 models the time-resolved case. This equation can be easily transform to frequency domain by Fourier transforming,

$$\left[ \frac{i\omega}{c} + \mu_a \right] \hat{\phi}(\mathbf{x}, \omega) - \nabla \cdot (D \nabla \hat{\phi}(\mathbf{x}, \omega)) = \hat{S}(\mathbf{x}, \omega), \quad (2.26)$$

where  $\omega$  is the angular frequency,  $\hat{\phi}$  and  $\hat{S}$  represent the Fourier transform of  $\phi$  and  $S$  respectively,  $i$  is the imaginary number and  $\mathbf{x}_0$  is the position of the source.

For the continuous-wave case, the Equation 2.24 simplifies to,

$$-\nabla \cdot (D(\mathbf{x}) \nabla \phi(\mathbf{x})) + \mu_a(\mathbf{x}) \phi(\mathbf{x}) = S(\mathbf{x}). \quad (2.27)$$

### 2.3. Diffuse optics systems and instrumentation

Diffuse optics spectroscopy and imaging is a promising biophotonics research field. Although, first works were published in the seventies, the foundations of the field were built in the late eighties, nineties and the beginning of this century. During the last twenty years, technology has evolved in order to improve the sensitivity to (1) deeper layers of the tissue and (2) absorption changes. In the following sections different technologies will be described and their main advantages and drawbacks will be assessed.

The main difference between diffuse optics technologies is based on the type of light source used. Continuous-wave (CW) based technology was the first system to be used in the clinics; its main characteristic is that the intensity of the source of light is constant over time and only the decrease of such intensity is measured (see Figure 2.6 left). By using the modified Beer-Lambert law [57] blood oxygenation can be estimated by near-infrared (NIR)

spectroscopy. Nevertheless, single source-detector distance CW technology suffers from two disadvantages. First, absorption and scattering cannot be decoupled. Second, superficial and deep layers cannot be distinguished from each other. Although multi-distance systems can overcome these problems by encoding depth using several source-detector distances. This technology could not be feasible in some clinical environments where maximum source-detector separation distances are constrained.

In order to overcome the first drawback, frequency domain (FD) technology appeared on the scene (see Figure 2.6 center). FD light source is frequency modulated usually at MHz order. Then, detectors measure not only the attenuation but also the phase change. This technology permits to decouple absorption from scattering in the inverse problem.

In the late eighties and early nineties, first time-domain or time-resolved (TR) system were developed [36, 57, 46]. In this case, the source is a pulse of light that can be mathematically modeled as a Dirac delta pulse (see Figure 2.6 right). The idea is to measure the photons distribution time-of-flight (DTOF), since the photons arrival times can be probabilistically associated with the penetration depth in the media.

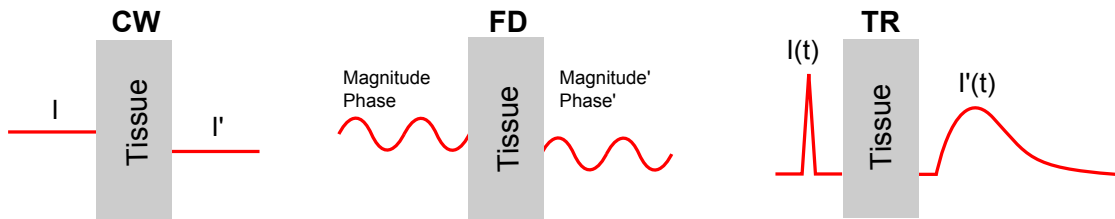


Figure 2.6: Diffuse optics systems. (Left) Continuous-wave system; only attenuation of light is measured,  $I' < I$ . (Middle) Frequency domain system; attenuation and dephasing are measured. (Right) Time-resolved system; the time of flight distribution of the arriving photons are measured. Image was inspired by [79].

| Technology | Uniqueness (1) | Cost (2) | Clinics (3) |
|------------|----------------|----------|-------------|
| CW         | No             | Low      | Yes         |
| FD         | Yes            | Medium   | Yes         |
| TR         | Yes            | High     | Starting    |

Table 2.2: Comparison of state-of-the-art diffuse optical technologies. (1) Uniqueness of the solution based on [6] and assuming refraction index is known; (2) relative comparison between diffuse optical technologies; (3) based on the daily basis clinical use and available industrial products.

### 2.3.1. Light sources

Diffuse optical systems do not require polarized light sources or high levels of directionality, since large scattering mediums make these properties go quickly to zero. Coherence is neither a useful property in diffuse optics systems. The most important properties of light sources at diffuse optics systems are: wavelength bandwidths, temporal profile and intensity.

For diffuse spectroscopy, the number of wavelengths used should be at least as many as the number of chromophores to be measured (this property follows from the Beer-Lambert law

and linear algebra theory). The bandwidth of sources does not need to be monochromatic because of the broad extinction spectrum of chromophores. Therefore, laser diodes with few nanometers bandwidth fulfill by far the requirements of diffuse spectroscopy systems [31, 156]. Even LED diodes that have spectral bandwidths of tens nanometers can still be used in practice [249, 253]. Lasers such as titanium-doped sapphire (Ti:sapphire) that can be tuned in 650-1100 nm range are also suitable [177, 253, 19]. Recently, supercontinuum lasers with wavelength selection systems have also been used [257].

The temporal profile is also a critical aspect of the light source. For TR systems light source must be pulsed at a frequency of maximum 80 MHz, that is, 12.5 ns of separation between light peaks to guarantee that no photon overlapping while occur between photon arrival profiles. Ti:sapphire laser is suited for time-resolved systems since it is intrinsically pulsed.

### 2.3.2. Optical detectors

Although there are many types of optical detectors, all of them are based on the same concept: convert photons into electrons and measure the electrical current that is generated. Detectors can be divided in two types based on the physical process that takes place. The first type uses photon induced emission of electrons from a photocathode; basically photons hit a photocathode and some of those photons will collide with electrons transferring to them their energy. The hit electrons will gain kinematic energy which will be measured by the detectors. Examples of such detectors are streak cameras and photomultiplier tubes (PMT). PMTs are build with several dynodes to multiply the generations of electrons, see Figure 2.7. The second type of detectors are based on the excitation of electrons into the conduction band of a semiconductor. Solid-state photodiodes (PD), avalanche photodiodes (APD), single photon avalanche diodes (SPAD) and charge-coupled devices (CDD) are examples of this kind of detectors.

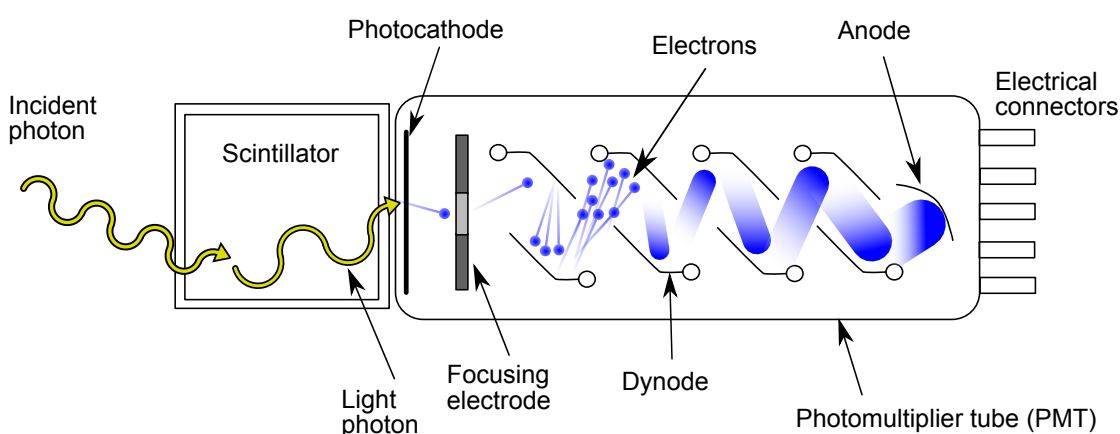


Figure 2.7: Scheme of a standard photomultiplier. Dynodes multiply the generation of electrons from a single photon. (Modified figure taken from Wikimedia Commons).

The main performance characteristics of a detector are the following:

- **Spectral sensitivity:** The sensitivity of a detector to a given wavelength depends mostly on the material that is used. For example, detectors such as PMT are suited

for 200-900 nm wavelength. However, PD and APD detectors detect deeper in near-infrared wavelengths than PMT. Moreover, CDD have a limitation up to 1000 nm [22].

- **Dynamic range and linearity:** To achieve quantitative results the response of the detector should be linear. Moreover, the dynamic range of detectors must be as high as possible because the light power decreases exponentially with distance. Usually dynamic ranges are expressed in decibels due to the large range of light intensity ratios a detector measures. The range where the detectors behaves linearly is usually known as the linear dynamic range.
- **Instrumental response function:** The physics and the electronics of the detector will determine how much time it will take since a photon hits the detector until voltage signal is received. This process is not instantaneous since photons have to be converted into electrons and their charge measured. This delay can be considered as random and depends on the type of detector used, see Figure 2.8.
- **Responsivity:** It is defined as the ratio between the incident radiant energy in the detector (watts,  $W$ ) and the generated current (amperes,  $A$ ). Responsivity is wavelength dependant and is usually defined as  $R = \eta e / (hf)$  where  $\eta$  is the quantum efficiency (conversion efficiency of photons into electrons),  $e$  is the electron electric charge,  $h$  is the Planck's constant and  $f$  is the frequency of the arriving photons. Units of responsivity are defined as the ration between amperes and watts.

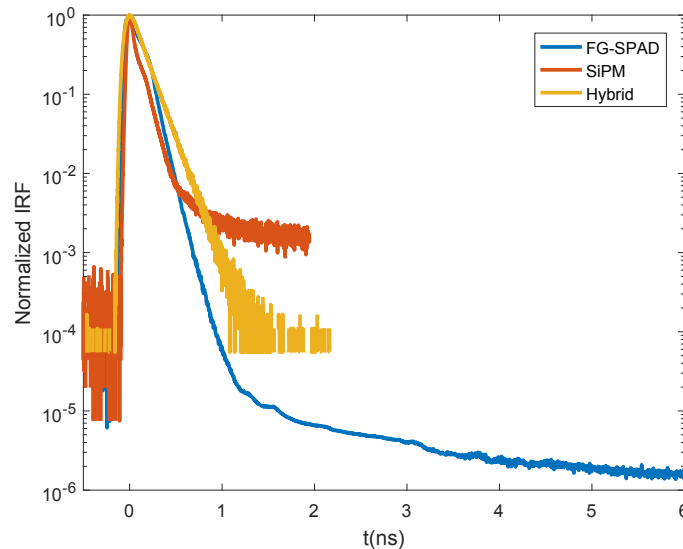


Figure 2.8: Instrumental response functions of several detector types. (FG-SPAD) Fast-gated SPAD. (SiPM) Silicon Photomultiplier. (Hybrid) SiPM and SPAD. These measurements were taken at Politecnico di Milano labs.

### 2.3.3. Noise

Classically, diffuse optics measurements are contaminated by several sources of noise. In the next sections I will describe the two main sources of these inaccuracies and their intrinsic nature.

### Shot noise

Shot noise, also known as photon noise, is an uncertainty related with the measurement of light. The number of photons reaching a detector is always an integer number because light is made up of photons which could be seen as discrete particles. Because of their discrete nature and the random fluctuations of arriving photons if the average number of detected photons is a decimal number the detector will not be able to express it and the detected value will fluctuate between integer numbers. This phenomenon is more evident when the intensity of light is low.

Shot noise depends on the signal magnitude and constitutes one of the biggest sources of noise when measuring light. Shot noise is not additive but is applied to the signal, that is, it depends on the intensity of the signal. As the number of detected photons must be an integer value the arrival of photons to a detector can be considered as a Poisson process. The modelling would be

$$Pr(N = k) = \frac{e^{-\lambda t} (\lambda t)^k}{k!}, \quad (2.28)$$

where  $\lambda$  is the number of arriving photons in a unit time interval and  $t$  is a time interval. Although for high light intensities noise can be modelled as Gaussian with intensity dependant variance. A widely known property of a Poisson distribution is that its expected value is equal to its variance. Therefore, the bigger the intensity of the signal the greater the noise in absolute terms. However, the signal-to-noise ratio (SNR) of the signal will increase with  $\sqrt{N}$  where  $N$  is the number of arriving photons. This shows that this noise plays a bigger role at low intensity levels, for example at late arrival photons in time-resolved systems. A conclusion taken from this is that in order to reduce the effect of shot noise, a higher intensity should be used; in fact if one wants the SNR to be increased by 10 then intensity must be increased by 100.

Shot noise can be interpreted as a lower bound of the noise a signal can have, that is, as shot noise is inherited from the quantum nature of photons that noise will also exist even if an ideal detector is used. Therefore, in the best case the signal will suffer from shot noise.

### Dark noise

Dark noise, also known as dark current, is a type of noise that happens even when there are no photon in the environment. It occurs because some electrons are thermally generated and are mistaken with the electrons generated by photoelectric effect. Dark noise is difficult to avoid and increases with the sensor time of exposure because it heats up. It also follows a Poisson distribution but usually is negligible in comparison with other types of noise.

#### 2.3.4. Time-resolved systems revisited

The main components of a time-resolved system are a pulsed laser and a detection system to capture the arrival time of each photon. There are two types of detection methods: streak cameras and time-correlated single-photon counting systems.

### Streak camera

Streak cameras provide the highest temporal resolution; for most commercial systems it is around 1-10 ps which is far enough for diffuse optical applications since latest photons arrive at the order of nanoseconds. The operating principle of streak cameras is based on a deflector that deviates entering photons depending on their arrival times and location. The result is a two-dimensional streak image where light intensity is mapped into space and time axes. A typical setup for a streak camera is shown at Figure 2.9. The light pulse is divided in three parts to inject it into the medium, to synchronize the streak camera with a sync signal and to set the laser pulse emission time ( $t_0$ ).

Although, streak cameras have the highest temporal resolution they are quite expensive and bulky, they suffer from low sensitivity and their dynamic range is not sufficient in some situations.



Figure 2.9: Standard setup for time-resolved systems based on streak cameras.

### Time-correlated single photon counting

Since streak cameras are bulky they are not suitable for clinical environments. As an alternative, time-correlated single photon counting (TCSPC) technology have been proposed. TCSPC technology allows to use PMT, hybrid-PMT or SPAD detectors which are smaller and cheaper than streak cameras. The use of Silicon photomultipliers (SiPM) detectors has also been reported [139, 185, 152]. Moreover, TCSPC cards can be easily integrated in a personal computer making it much easier to implement into a system.

The idea behind TCSPC is to send light pulses with attenuated intensity so that at most one photon reaches the detector for each pulse. Each photon arrival time will be recorded and after several millions of pulses a histogram will be created. The histogram represents the DTOF. The main difference between streak cameras and TCSPC is that the former measures a DTOF for each pulse and adds up all of them and the latter detects at most the arrival time of one photon for each pulse and finally builds a DTOF histogram from all of recorded times.

If more than one photon from the same pulse arrives to the detector, only the first photon will be detected. In that case, early photons will be over-represented and the DTOF histogram will not be correct. For that reason, it is important to impose that at most one photon reaches a detector for each pulse. Two conditions are needed to make this possible. First, laser power must be adjusted (that is, the number of photons launched must be low) so that a detector counts a photon only for 1 to 10% of the laser pulses. Second, the pulse frequency should be low enough to (1) prevent photon overlapping and to (2) wait for detectors dead-time. Overlapping takes place when a photon arrives to the detector after the next pulse has already been flashed. Detector dead-time is the time it takes to electronics to recover from



a photon detection [230]. An 80 MHz repetition frequency is usually far enough to avoid overlapping.

A typical TCSPC based time-resolved system is shown at Figure 2.10. The level trigger receives the sync signal from the laser pulse to record when the pulse was flashed. The Constant Fraction Discriminator (CFD) computes the photon arrival time. In this case a classic level trigger could not be used because for some detectors, such as PMT, the amplitude of the detected voltage changes at each detection (this happens due to the physics of PMT. When a photon hits the detector a bunch of electrons are generated whose quantity varies from detection to detection). Therefore, CFD discriminates noise and adjusts photon arrival time depending on these factors. The Time to Amplitude Converter (TAC) transforms the time difference between the pulse and photon arrival time to a voltage. Then, the Analog to Digital Converter (ADC) retrieves the digital timing value of the photon arrival time. Finally, a field-programmable gate array will compute the histogram with all timing values.

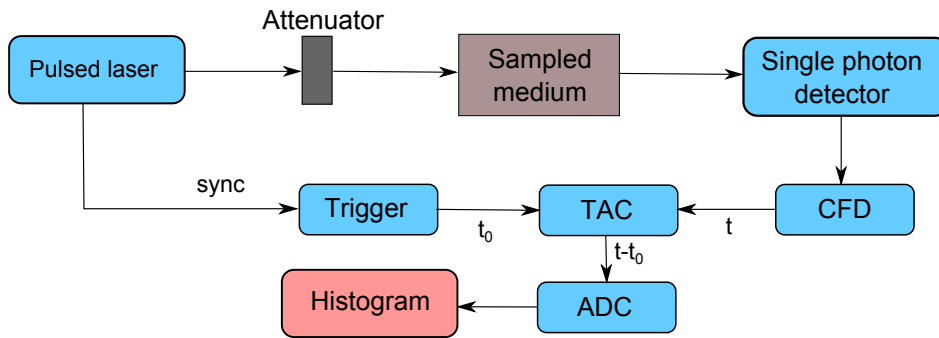


Figure 2.10: Standard setup for time-resolved systems based on TCSPC technology. CFD stands for Constant Fraction Discriminator, TAC for Time to Amplitude converter and ADC for Analog to Digital Converter.

## 2.4. Numerical methods for photon propagation models

As models described at Subsection 2.2.2 (see Equations 2.23 and 2.24) can be solved numerically using a computer, there are several numerical methods to solve them. Monte-Carlo method is usually used to solve the Radiative Transfer equation and the Finite Element method for the Diffusion equation. In the following subsections a technical description of both methods is given and some implementation remarks are pointed out.

### 2.4.1. Monte-Carlo method

Let us assume that a Monte-Carlo simulation is performed in a  $N$ -layer model with arbitrary shape. For simplicity it will be assumed that there are no void regions, although it would be straightforward to include that possibility in the method. Launching  $M$  photons, the workflow will be the following:

1. **Starting point:** Each photon is launched from the source point with the direction to which the light source is pointing (usually perpendicular to the boundary).

2. Each photon will travel a distance of

$$d = -\frac{\ln \varepsilon}{\mu_s}, \quad (2.29)$$

where  $\varepsilon \in \text{Unif}[0, 1]$  is a uniformly distributed random number. For each photon a different pseudo-random number will be generated. See derivation of Equation 2.29 at Appendix B.

3. Check if the photon has left the domain. If yes, then update the distance that it has travelled in each layer and stop simulating that photon, if not then go to next step.
4. After each photon has travelled a distance then it will scatter to a direction. To determine the direction of scattering, the phase function and anisotropic factor have to be used. As was stated previously, the Henyey-Greenstein phase function with an anisotropic factor of  $g = 0.9$  is a good approximation for human tissue. Therefore, to calculate a random number that follows Henyey-Greenstein distribution the following equation has to be used (see Derivation at Appendix B),

$$\cos \theta = \frac{1}{2g} \left[ 1 + g^2 - \left( \frac{1 - g^2}{1 + g\varepsilon} \right) \right], \quad (2.30)$$

where  $\varepsilon \in \text{Unif}[0, 1]$  is a uniformly distributed random number. The azimuth angle follows,

$$\varphi = 2\pi\varepsilon, \quad (2.31)$$

since it is assumed to be isotropic in that coordinate.

5. For each photon, update the distance that it has travelled in each layer.
6. Go to step 2 until all photons have left the computational domain.
7. **Ending point:** Calculate  $M$  uniformly distributed random numbers,  $U_i \in \text{Unif}[0, 1]$ . For each photon calculate

$$W_i = 1 - \sum_{j=1}^N e^{-\mu_{a,j} L_{i,j}} \quad (2.32)$$

where  $i$  is the index of each photon,  $j$  is the index of each layer,  $\mu_{a,j}$  indicates the absorption at layer  $j$  and  $L_{i,j}$  is the distance travelled by photon  $i$  at layer  $j$ . If the random number  $U_i$  is greater than  $W_i$  then photon  $i$  survives, if not it is assumed that it was absorbed. Finally, the simulated DTOF histogram for each detector will be generated using the photons that survived.

Although there are several ways of implementing the Monte-Carlo method for photon propagation, the method explained here has the following advantages:

- Since photons are simulated until they leave the domain and for each of them, the distance travelled at each layer is recorded, then the simulation is absorption independent. In other words, the effect of absorption is computed after the simulation, so if the absorption distribution is changed the simulation has not to be redone. This property does not hold for scattering changes [234].
- It is easy to parallelize because each photon is assumed to be independent and they do not have interaction with each other [69, 4].

The open-source code MMCLAB [68] follows the scheme described here. MMCLAB models the domain using a tetrahedral mesh so it can be applied to complex structures like a human head or limb. The MCX version also supports GPU computing and can be run in MATLAB or GNU Octave [151].

### 2.4.2. Finite Element method (FEM)

First, a detailed derivation of Finite Element Method (FEM) for Diffusion Approximation equation is given. The weak formulation of the equation is derived and its associated linear systems in time, frequency and continuous-wave versions are presented. After, computational and implementation details of FEM are discussed.

#### The FEM formulation of the Diffusion approximation

Let us define the time-resolved Diffusion Approximation problem as:

$$\begin{cases} \frac{1}{c} \frac{\partial \phi(\mathbf{x}, t)}{\partial t} - \nabla \cdot (D(\mathbf{x}) \nabla \phi(\mathbf{x}, t)) + \mu_a(\mathbf{x}) \phi(\mathbf{x}, t) = S(\mathbf{x}, t), & \mathbf{x} \in \Omega, \quad t > 0 \\ \phi(\mathbf{x}, t) - 2C_{RD} \frac{\partial \phi(\mathbf{x}, t)}{\partial n} = 0, & \mathbf{x} \in \partial\Omega, \quad t > 0 \\ \phi(\mathbf{x}, t) = 0, & \mathbf{x} \in \Omega, \quad t = 0 \end{cases} \quad (2.33)$$

where  $S(\mathbf{x}, t)$  is an arbitrary source function,  $\Omega$  represents the space domain and  $\partial\Omega$  the boundary.

The solution of the PDE defined at Equation 2.33 must be in the space  $C^k(\Omega)$  defined as

$$C^k(\Omega) = \left\{ u \in C(\Omega) \cap \frac{\partial^{i+j} u}{\partial^i x \partial^j y} \in C(\Omega) \right\}, \quad (2.34)$$

where  $i, j \geq 0$ ,  $i + j \leq k$  and  $C(\Omega)$  is a continuous function at domain  $\Omega$ . The equality of the PDE still holds if both parts are multiplied by an arbitrary function  $\psi$  and integrated over the whole domain,

$$\int_{\Omega} \frac{1}{c} \frac{\partial \phi}{\partial t} \psi - \nabla \cdot (D \nabla \phi) \psi + \mu_a \phi \psi \, d\mathbf{x} = \int_{\Omega} S \psi \, d\mathbf{x}. \quad (2.35)$$

If  $\phi$  holds for a particular  $\psi$ , it cannot be said that it also holds for Equation 2.33. Nevertheless, let us define a ball of radius  $\delta$ ,

$$B_{\delta}(\mathbf{x}') = \{ \mathbf{x} \in \Omega \mid \|\mathbf{x} - \mathbf{x}'\|_2 < \delta \}, \quad (2.36)$$

where  $\mathbf{x}'$  is the center of the ball and  $\psi \in C^2(\Omega)$  follows the properties:

- $\psi(\mathbf{x}) > 0$  if  $\mathbf{x} \in B_{\delta}$ ,
- $\psi(\mathbf{x}) = 0$  if  $\mathbf{x} \notin B_{\delta}$ ,
- $\int \psi \, d\mathbf{x} = 1$ .

When  $\delta \rightarrow 0$ , the  $\psi$  function will converge to a Dirac delta distribution which implies that Equation 2.35 must be true at point  $(\mathbf{x}', t)$  for  $t \in [0, \infty)$ . Imposing that Equation 2.35 must be fulfilled for any function  $\psi \in C^k(\Omega)$ , implies that Equation 2.35 will be fulfilled at every point  $(\mathbf{x}', t)$  which is equivalent to state that Equation 2.35 has the same solution as the boundary value problem (BVP) of Equation 2.33.

The chain rule  $\nabla \cdot (D\nabla\phi \psi) = \nabla \cdot (D\nabla\phi)\psi + D\nabla\phi \cdot \nabla\psi$  and the Gauss/divergence theorem (see Appendix C) converts Equation 2.35 to:

$$\frac{1}{c} \int_{\Omega} \frac{\partial\phi}{\partial t} \psi \, d\Omega - \int_{\partial\Omega} D \frac{\partial\phi}{\partial n} \psi \, d\Omega + \int_{\Omega} D\nabla\phi \cdot \nabla\psi \, d\Omega + \int_{\Omega} \mu_a \phi \psi \, d\Omega = \int_{\Omega} S\psi \, d\Omega. \quad (2.37)$$

Applying the boundary conditions the following equation holds,

$$\frac{1}{c} \int_{\Omega} \frac{\partial\phi}{\partial t} \psi \, d\Omega - \int_{\partial\Omega} \frac{\phi}{2C_R} \psi \, d\Omega + \int_{\Omega} D\nabla\phi \nabla\psi \, d\Omega + \int_{\Omega} \mu_a \phi \psi \, d\Omega = \int_{\Omega} S\psi \, d\Omega. \quad (2.38)$$

Looking at the equation above it is clear that the function requirements are  $\psi \in H^1(\Omega)$  and  $\phi \in \mathcal{S}$ , where

$$\begin{aligned} L^2(\Omega) &= \left\{ \psi : \int_{\Omega} \psi^2 \, d\Omega < \infty \right\} \\ H^1(\Omega) &= \left\{ \psi \in L^2(\Omega) : \frac{\partial\psi}{\partial x}, \frac{\partial\psi}{\partial y} \in L^2(\Omega) \right\}, \\ \mathcal{S}([0, T]) &= \left\{ \phi : \phi(\cdot, t) \in H^1(\Omega), t \in [0, T] \right\}, \end{aligned} \quad (2.39)$$

Taking into account all aspects described previously, the *weak formulation* of the BVP at Equation 2.33 is

For any  $t \in [0, T]$  find  $\phi \in \mathcal{S}$  such that

$$\begin{aligned} \frac{1}{c} \int_{\Omega} \frac{\partial\phi}{\partial t} \psi \, d\Omega - \int_{\partial\Omega} \frac{\phi}{2C_R} \psi \, d\Omega + \int_{\Omega} D\nabla\phi \nabla\psi \, d\Omega + \int_{\Omega} \mu_a \phi \psi \, d\Omega &= \int_{\Omega} S\psi \, d\Omega, \\ \forall \psi \in H^1(\Omega). \end{aligned} \quad (2.40)$$

Since Equation 2.40 must be fulfilled for all  $\psi \in H^1(\Omega)$  then it is assured that the BVP is equivalent to the weak formulation. This formulation is called *weak* since less restrictions are posed to the solution compared to the original BVP; note how the solution at BVP should be twice derivable but at the weak formulation only squared integrability is assumed.

The weak formulation can be rewritten as

For any  $t \in [0, T]$  find  $\phi \in \mathcal{S}$  such that

$$a(\phi, \psi) = l(\psi), \quad \forall \psi \in H^1(\Omega) \quad (2.41)$$

where

$$a(\phi, \psi) = \frac{1}{c} \int_{\Omega} \frac{\partial \phi}{\partial t} \psi \, d\Omega - \int_{\partial\Omega} \frac{\phi}{2C_R} \psi \, d\Omega + \int_{\Omega} D \nabla \phi \nabla \psi \, d\Omega + \int_{\Omega} \mu_a \phi \psi \, d\Omega, \quad (2.42)$$

$$l(\psi) = \int_{\Omega} S \psi \, d\Omega,$$

and  $a(\phi, \psi)$  is bilinear and  $l(\psi)$  linear.

Let us approximate  $V = H^1(\Omega)$  and  $W = \mathcal{S}$  spaces using the  $n$ -dimensional subspaces  $V_n \subset V$  and  $W_n \subset W$ . Now, problem formulation is:

For any  $t \in [0, T]$  find  $\phi_n \in W_n$  such that

$$a(\phi_n, \psi_n) = l(\psi_n), \quad \forall \psi_n \in V_n, \quad (2.43)$$

which is known as the *Galerkin formulation*.

The Galerkin orthogonality principle (see Section 3 of [82]) ensures that  $\phi_n$  is the best approximation of the solution  $\phi$  at space  $V_n$ .

Let  $\{\varphi_i\}_{i=1}^n$  be a basis of the subspace  $V_n$ . Equation 2.43 is satisfied for all  $\psi_n \in V_n$  if and only if it is satisfied for each basis  $\varphi_i$  (this can be proven by using the linearity of the equation and realizing that any  $\psi_n$  can be represented by a linear combination of the linear basis). Therefore, using the linear combination  $\psi_n = \sum_{i=1}^n \alpha_i \varphi_i$  where  $\alpha_i$  is a scalar, the problem of Equation 2.43 transforms to

$$\sum_{j=1}^n \alpha_j a(\phi_n, \varphi_j) = \sum_{j=1}^n \alpha_j l(\varphi_j), \quad (2.44)$$

which implies that equation is satisfied if

$$a(\phi_n, \varphi_j) = l(\varphi_j), \quad \forall j = 1, \dots, n. \quad (2.45)$$

Approximation to the solution  $\phi_n$  can also be represented using the linear combination  $\phi_n(\mathbf{x}, t) = \sum_{i=1}^n \Psi_i(t) \varphi_i(\mathbf{x})$  where the basis functions are not dependent on time. It follows that

$$\sum_{i=1}^n a(\Psi_i \varphi_i, \varphi_j) = l(\varphi_j), \quad \forall j = 1, \dots, n. \quad (2.46)$$

Equation 2.46 is a linear system where  $\Psi_j(t)$  are the unknowns. In the case of time-resolved Diffusion Approximation the linear system is:

$$\sum_{i=1}^n \frac{\partial \Psi_i}{\partial t} \frac{1}{c} \int_{\Omega} \varphi_i \varphi_j \, d\Omega - \frac{\Psi_i}{2C_R} \int_{\partial\Omega} \varphi_i \varphi_j \, d\Omega + \Psi_i \int_{\Omega} D \nabla \varphi_i \nabla \varphi_j \, d\Omega + \Psi_i \int_{\Omega} \mu_a \varphi_i \varphi_j \, d\Omega$$

$$= \int_{\Omega} S \varphi_j \, d\Omega, \quad \forall j = 1, \dots, n. \quad (2.47)$$

The above formulation can be described in matrix form by defining

$$\begin{aligned}
\mathbf{H} &= \frac{1}{c} \int_{\Omega} \varphi_i \varphi_j \, d\Omega \\
\mathbf{M} &= \int_{\Omega} D \nabla \varphi_i \nabla \varphi_j \, d\Omega \\
\mathbf{K} &= \int_{\Omega} \mu_a \varphi_i \varphi_j \, d\Omega \\
\mathbf{N} &= \frac{1}{2C_R} \int_{\partial\Omega} \varphi_i \varphi_j \, d\Omega \\
\mathbf{s} &= \int_{\Omega} S \varphi_j \, d\Omega
\end{aligned} \tag{2.48}$$

and the linear system of Equation 2.47 is equivalent to

$$\mathbf{H} \frac{\partial \Psi}{\partial t} + [\mathbf{M} + \mathbf{K} + \mathbf{N}] \Psi = \mathbf{s}, \tag{2.49}$$

where  $\Psi = [\Psi_1, \dots, \Psi_n]^T$ . Equation above is the most general form of the discretized version for the Diffusion Approximation. In the subsections below, the time (in)dependent and frequency cases are described separately.

### Time-independent case

When the source is constant over time, the time derivative is zero. Therefore, for the time-independent case, the linear system to solve is:

$$[\mathbf{M} + \mathbf{K} + \mathbf{N}] \Psi = \mathbf{s}. \tag{2.50}$$

It should be noted that the time-independent solution is equivalent to the time-dependent solution integrated in time,

$$\mathbf{H} \int_0^{\infty} \frac{\partial \Psi}{\partial t} \, dt + [\mathbf{M} + \mathbf{K} + \mathbf{N}] \int_0^{\infty} \Psi(t) \, dt = \int_0^{\infty} \mathbf{s}(t) \, dt. \tag{2.51}$$

The first term is equal to zero because

$$\int_0^{\infty} \frac{\partial \Psi}{\partial t} \, dt = \lim_{t \rightarrow \infty} \int_0^t \frac{\partial \Psi}{\partial t'} \, dt' = \lim_{t \rightarrow \infty} [\Psi(t)]_0^t = 0, \tag{2.52}$$

since  $\Psi_i(0) = 0$  and  $\Psi_i(t)$  tends to zero at infinity. The source term is a Dirac delta,

$$\int_0^{\infty} \mathbf{s}(t) \, dt = \lim_{t \rightarrow \infty} \int_0^t \mathbf{s}(t') \, dt' = \lim_{t \rightarrow \infty} \int_0^t \delta(\mathbf{x} - \mathbf{x}_s) \delta(t' - t_s) \, dt' = \delta(\mathbf{x} - \mathbf{x}_s). \tag{2.53}$$

Therefore, Equation (2.51) is

$$[\mathbf{M} + \mathbf{K} + \mathbf{N}] \int_0^{\infty} \Psi(t) \, dt = \delta(\mathbf{x} - \mathbf{x}_s), \tag{2.54}$$

and  $\int_0^{\infty} \Psi(t) \, dt$  is the solution for continuous-wave case.

### Time-dependent case

The time derivative appearing in Equation (2.49) has to be discretized. A good option is to use the  $\theta$ -scheme with  $\theta = 1/2$ , also known as *Crank-Nicolson* scheme. Applying the  $\theta$ -scheme to Equation (2.49) it holds,

$$\frac{\Psi^{k+1} - \Psi^k}{\Delta t} \mathbf{H} + \theta \Psi^{k+1} [\mathbf{M} + \mathbf{K} + \mathbf{N}] + (1 - \theta) \Psi^k [\mathbf{M} + \mathbf{K} + \mathbf{N}] = \theta \Psi^{k+1} \mathbf{s} + (1 - \theta) \Psi^k \mathbf{s}. \quad (2.55)$$

The  $\theta = 0$  case stands for an *explicit scheme* and the  $\theta = 1$  case for an *implicit scheme*. For the Crank-Nicolson case ( $\theta = 1/2$ ) the following system should be solved,

$$\left[ \frac{1}{\Delta t} \mathbf{H} + \frac{1}{2} (\mathbf{M} + \mathbf{K} + \mathbf{N}) \right] \Psi^{k+1} + \left[ \frac{1}{2} (\mathbf{M} + \mathbf{K} + \mathbf{N}) - \frac{1}{\Delta t} \mathbf{H} \right] \Psi^k = \frac{\mathbf{s}^{k+1} + \mathbf{s}^k}{2} = \frac{\mathbf{s}^k}{2}, \quad (2.56)$$

where the term  $\mathbf{s}^{k+1} = 0$ , because the short time pulse is only nonzero in the first time interval. Therefore, taking into account that  $\Psi^0 = 0$ , because at the beginning of the experiment there is no light, the first time step is,

$$\mathbf{U} \Psi^1 = \frac{\mathbf{s}^0}{2}, \quad k = 0 \quad (2.57)$$

and in the following

$$\mathbf{U} \Psi^{k+1} = \mathbf{V} \Psi^k, \quad \forall k \geq 1, \quad (2.58)$$

where  $\mathbf{U} = \frac{1}{\Delta t} \mathbf{H} + \frac{1}{2} (\mathbf{M} + \mathbf{K} + \mathbf{N})$  and  $\mathbf{V} = \frac{1}{\Delta t} \mathbf{H} - \frac{1}{2} (\mathbf{M} + \mathbf{K} + \mathbf{N})$ .

### Frequency-dependent case

After applying the Fourier transform to the time-dependant case, the linear system to solve is

$$\left[ \frac{i\omega}{c} \mathbf{H} + \mathbf{M} + \mathbf{K} + \mathbf{N} \right] \hat{\Psi} = \hat{\mathbf{s}}, \quad (2.59)$$

where  $\hat{\Psi}$  is the Fourier transform of  $\Psi$  and  $\omega$  the angular frequency.

### 2.4.3. Computational and implementation details of FEM

Now, once the mathematical theory behind FEM has being described, I will give the nitty-gritty details of its computational implementation.

Let us define a two-dimensional irregular mesh with  $M$  elements and  $N$  nodes. Most used two-dimensional elements are triangles and rectangles, in Figure 2.11 left an example of a triangular mesh in a circular domain is shown. A limited support basis function is assigned to each node to minimize the number of nonzero elements in the matrices, (see Figure 2.11 right). The linear basis functions for each node are defined as

$$\varphi_i(\mathbf{x}_j) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}, \quad (2.60)$$

where  $\mathbf{x}_j$  is the position of node  $j$ . The polynomials  $\varphi_i$  can be of any order, the larger the order the better the approximation.

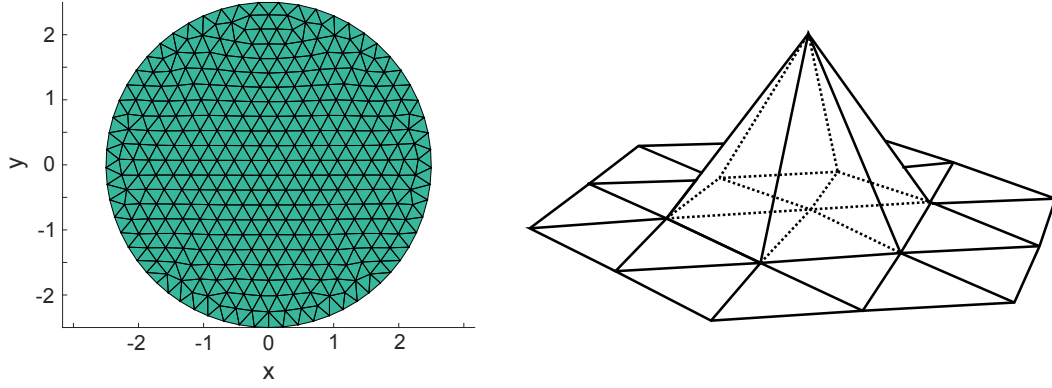


Figure 2.11: (Left) Two-dimensional triangular mesh of a circumference. (Right) Linear basis function associated to a mesh node.

### Reference triangle

Let us examine the case where two basis function are associated to nodes in the same element, (see Figure 2.12). For an arbitrary triangle  $\tau$  with vertices at

$$\begin{cases} \mathbf{v}_1 = (x_1, y_1) & (2.61) \\ \mathbf{v}_2 = (x_2, y_2) & (2.62) \\ \mathbf{v}_3 = (x_3, y_3) & (2.63) \end{cases}$$

there are three basis function  $\phi_1$ ,  $\phi_2$  and  $\phi_3$  associated to each node. One way to compute  $\int \phi_1 \phi_2 dx$  is to do it directly in the arbitrary triangle, nevertheless this method will get more complicated for higher dimensions and standard numerical integration techniques must be adapted accordingly. For these reasons, it is better to compute the integral in a reference triangle  $\tau'$  with nodes at

$$\begin{cases} \mathbf{v}'_1 = (1, 0) & (2.64) \\ \mathbf{v}'_2 = (0, 1) & (2.65) \\ \mathbf{v}'_3 = (0, 0). & (2.66) \end{cases}$$

The explicit form of an affine transformation  $\mathcal{F} : \tau' \rightarrow \tau$  that maps the reference triangle into an arbitrary triangle element  $\tau$  is,

$$\mathcal{F}(\mathbf{x}') = \mathbf{A}^T \mathbf{x}' + \mathbf{b}, \quad (2.67)$$

where

$$\mathbf{A} = \begin{bmatrix} x_1 - x_3 & y_1 - y_3 \\ x_2 - x_3 & y_2 - y_3 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} x_3 \\ y_3 \end{bmatrix} \quad (2.68)$$

and  $\mathbf{x}'$  is a point inside the reference triangle.

For a function  $u$  defined in the arbitrary triangle and its equivalent function  $u'$  in the reference triangle, the following relationship holds:

$$u'(\mathbf{x}') = u(\mathcal{F}(\mathbf{x}')), \quad (2.69)$$



whose gradient is

$$\nabla' u' = \mathbf{A} \nabla u. \quad (2.70)$$

The equivalence of the integrals are,

$$\begin{aligned} \int_{\tau} \phi_1 \phi_2 \, d\mathbf{x} &= \int_{\tau'} \phi'_1 \phi'_2 |\det(\mathbf{A})| \, d\mathbf{x}', \\ \int_{\tau} \nabla \phi_1 \cdot \nabla \phi_2 \, d\mathbf{x} &= \int_{\tau'} (\mathbf{A}^{-1} \nabla' \phi'_1) \cdot (\mathbf{A}^{-1} \nabla' \phi'_2) |\det(\mathbf{A})| \, d\mathbf{x}', \end{aligned} \quad (2.71)$$

where  $d\mathbf{x} = |\det(\mathbf{A})| d\mathbf{x}'$  due to the change of variables.

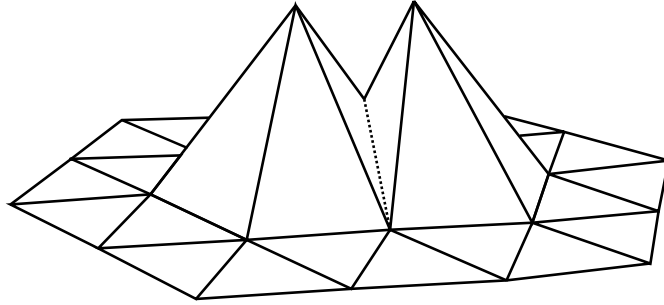


Figure 2.12: Graphical representation of two overlapping basis functions. The integral obtained from the multiplication of those two basis functions is only non-zero when the basis functions overlap; for linear basis functions, this only occur when they are adjacent.

### Gaussian quadrature

Integrals need to be computed numerically, since absorption and diffusion coefficient are just defined at node positions. Gaussian quadrature is one of the most popular numerical integration techniques for triangular and tetrahedral meshes. Gaussian quadrature approximates the integral of a function  $f(x, y)$  over a reference triangle with a discrete weighted summation,

$$\int_{\tau'} f(x, y) \, d\mathbf{x} \approx \sum_{i=1}^n w_i f(x_i, y_i), \quad (2.72)$$

where  $(x_i, y_i)$  are selected points in the reference triangle and  $w_i$  are the weights associated to those points. The goal is to obtain a set of positions and weights that can approximate the integral as accurately as possible. Discrete points have to be symmetrical in order to promote stability of the method. Commonly, quadrature rules are assigned a degree  $N$  to represent up to what polynomial order the numerical integration is exact. That is, Gaussian quadrature defined at Equation 2.72 is said to be exact up to degree  $N$  when

$$\int_{\tau'} P_N(x, y) \, d\mathbf{x} = \sum_{i=1}^{\mathcal{D}(N)} w_i P_N(x_i, y_i), \quad (2.73)$$

where  $P_N$  is a polynomial up to order  $N$  and  $\mathcal{D}(N)$  is the number of discrete points for the quadrature rule of order  $N$ .

The first order Gaussian quadrature for triangles must be exact for polynomials  $f(x, y) = 1$ ,  $f(x, y) = x$  and  $f(x, y) = y$ . A practical identity to compute Gaussian quadrature points is:

$$\int_{\tau'} x^i y^j d\mathbf{x} = \frac{i!j!}{(i+j+2)!}. \quad (2.74)$$

That is, using previous identity for first degree polynomials the following equalities must hold:

$$\frac{1}{2} = \sum_{i=1}^{\mathcal{D}(1)} w_i, \quad \frac{1}{6} = \sum_{i=1}^{\mathcal{D}(1)} w_i x_i, \quad \text{and} \quad \frac{1}{6} = \sum_{i=1}^{\mathcal{D}(1)} w_i y_i. \quad (2.75)$$

In this case, it is clear that  $\mathcal{D}(1) = 1$  and  $w_1 = 1/2$ ,  $x_i = 1/3$  and  $y_i = 1/3$ . Therefore, since the only discrete point is at the barycentre, the quadrature is symmetric in this case.

For second degree polynomials ( $f(x, y) = \{1, x, y, xy, x^2, y^2\}$ ), the equalities that must hold are:

$$\begin{aligned} \frac{1}{2} &= \sum_{i=1}^{\mathcal{D}(2)} w_i, & \frac{1}{6} &= \sum_{i=1}^{\mathcal{D}(2)} w_i x_i, & \frac{1}{6} &= \sum_{i=1}^{\mathcal{D}(2)} w_i y_i, \\ \frac{1}{24} &= \sum_{i=1}^{\mathcal{D}(2)} w_i x_i y_i, & \frac{1}{12} &= \sum_{i=1}^{\mathcal{D}(2)} w_i x_i^2, & \frac{1}{12} &= \sum_{i=1}^{\mathcal{D}(2)} w_i y_i^2. \end{aligned} \quad (2.76)$$

There exists a solution for  $\mathcal{D}(2) = 2$ , however the discrete points will not be symmetric, which is not a good property. For this reason,  $\mathcal{D}(2) = 3$  will be used whose discrete points are  $(x_1, y_1) = (1/6, 1/6)$ ,  $(x_2, y_2) = (2/3, 1/6)$  and  $(x_3, y_3) = (1/6, 2/3)$ . Associated weights are  $w_1 = w_2 = w_3 = 1/6$ .

For the following orders, the same procedure must be followed. In [62] exact Gaussian quadrature rules up to order twenty are presented. For tetrahedral geometries [251, 248] present techniques based on the same concepts explained above.

### Sparsity and symmetry

One of the main advantages of FEM is that matrices of the linear system are sparse and symmetric. This follows because basis function have a limited support and for most  $(i, j)$  pairs the expressions  $\varphi_i \varphi_j$  and  $\nabla \varphi_i \nabla \varphi_j$  will be null at all the domain. At Figure 2.13 the sparsity pattern of matrices  $\mathbf{H}$  and  $\mathbf{N}$  are shown.

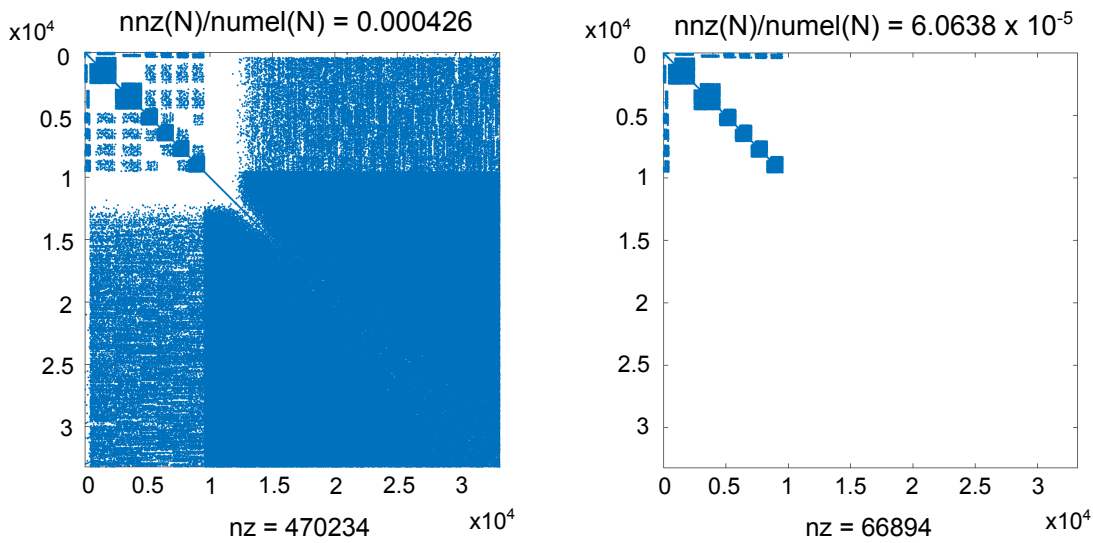


Figure 2.13: (Left) Sparsity pattern of  $\mathbf{H}$  matrix; note that matrices  $\mathbf{M}$  and  $\mathbf{K}$  have the same pattern. Due to the large number of nodes that the mesh has the sparsity cannot be fully appreciated in the image; nevertheless, non-zero values percentage is around 0.04%. (Right) Sparsity pattern of  $\mathbf{N}$  matrix, the matrix that implements the boundary conditions. Non-zero values percentage is around 0.006%.

Most of these matrices are maintained constant over all computations. For example, at Equation 2.58 the matrix  $\mathbf{U}$  is constant over all time steps. Therefore, it may be factorized to solve the linear systems much faster. Same property holds for standard and Mellin–Laplace datatypes that will be presented in the next section; the matrix of the system is always constant and by factorizing it the process can be made much faster. In practice, *LU*-factorization [83] is one of the most used methods and customized algorithms have been developed for large sparse matrices [30]. *LU* decomposition maintains a good ratio between the amount of saved time to solve linear systems and its computing time complexity.

## 2.5. Diffuse optical tomography

During the early development of diffuse optics, most of the research was focused on quantifying several physiological parameters such as blood oxygenation, saturation or flow volume using just a few sources and detectors. When the understanding of the physical principles behind diffuse optics were better, the computation capacity was improved and the instrumentation cost reduced researchers started to develop mathematical methods for computing two and three dimensional maps of optical and physiological properties. The basic idea is to recover the optical properties inside a volumetric medium by injecting light from the surface and collecting backscattered photons with detectors located at different positions. The measured photon arrivals are introduced in a tomographic algorithm and three-dimensional optical properties reconstruction is performed (see Figure 2.14). This approach is known as diffuse optical tomography (DOT).

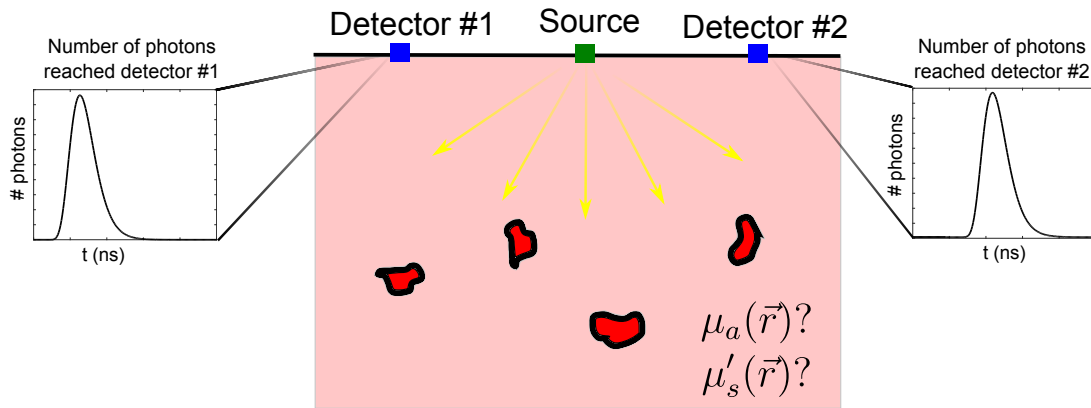


Figure 2.14: Typical setup of a time-resolved diffuse optical tomography problem. Sources and detectors are located at the surface of the tissue. Light is injected in the medium and some photons will scatter until they reach a detector. The diffuse medium will have some absorption and scattering ( $\mu_a, \mu'_s$ ) inhomogeneities.

Mathematically, DOT can be posed as an inverse problem. That is, from the photons backscattered to the detectors, one wants to retrieve the optical properties inside the medium; this is the opposite of the forward process (see Figure 2.15).

Inverse problems are not well-posed in terms of Hadamard properties, that is, a solution may not exist, it could be non-unique and may differ significantly for small changes in the initial conditions. Moreover, when such problems are discretized and linearized, the resulting systems are highly ill-conditioned and underdetermined. Due to these reasons regularization is usually needed in order to obtain physically feasible results.

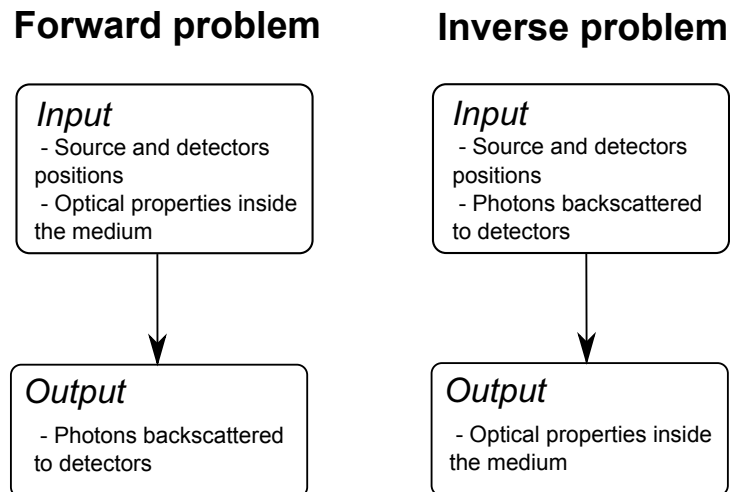


Figure 2.15: Comparison of the forward and inverse problems. In the first case the developed model will try to simulate the propagation of photons in a diffuse medium. For the latter an inverse model recovers the optical properties of the medium given the data collected at detectors.

In the following subsections, I will introduce the mathematical basics to solve the DOT inverse problem, derive the Born and Rytov approximations and sketch the structure of

a reconstruction algorithm. After, I will describe the reconstructions techniques used for continuous-wave, frequency- and time-resolved technologies.

### 2.5.1. Basics of optical properties reconstruction

In this section, I will focus on the time-resolved case since it can be seen as a generalization of other technologies. However, results can be easily derived for continuous-wave and frequency-resolved technologies.

Let us define a non-homogeneous linear PDE,

$$L\phi(\mathbf{x}, t) = f(\mathbf{x}, t), \quad (2.77)$$

where  $L$  is a linear operator,  $f(\mathbf{x}, t)$  is the source function and  $\phi(\mathbf{x}, t)$  is the unknown solution.

The Green function is defined as the solution to,

$$LG(\mathbf{x}, \mathbf{x}_s; t, t_s) = \delta(\mathbf{x} - \mathbf{x}_s; t - t_s), \quad (2.78)$$

where  $G(\mathbf{x}, \mathbf{x}_s; t, t_s)$  is the Green function for a Dirac delta distribution located at  $\mathbf{x}_s$  that burst at time  $t_s$  which is defined as,

$$\delta(\mathbf{x} - \mathbf{x}_s; t - t_s) = \begin{cases} \infty, & \text{if } \mathbf{x}_s = \mathbf{x} \cap t = t_s \\ 0, & \text{if } \mathbf{x}_s \neq \mathbf{x} \cup t \neq t_s \end{cases}, \quad \int_0^T \int_{\Omega} \delta(\mathbf{x} - \mathbf{x}_s; t - t_s) \, d\mathbf{x} dt = 1, \quad (2.79)$$

assuming  $T > t_s$ .

Therefore,

$$\int_0^T \int_{\Omega} LG(\mathbf{x}, \mathbf{x}'; t, t') f(\mathbf{x}', t') \, d\mathbf{x}' dt' = \int_0^T \int_{\Omega} \delta(\mathbf{x} - \mathbf{x}'; t - t') f(\mathbf{x}', t') \, d\mathbf{x}' dt' = f(\mathbf{x}, t), \quad (2.80)$$

and because  $L$  operator only applies to variables  $\mathbf{x}$  and  $t$  it follows,

$$L \int_0^T \int_{\Omega} G(\mathbf{x}, \mathbf{x}'; t, t') f(\mathbf{x}', t') \, d\mathbf{x}' dt' = f(\mathbf{x}, t). \quad (2.81)$$

The last equation implies that the solution to Equation 2.77 can be expressed as,

$$\phi(\mathbf{x}, t) = \int_0^T \int_{\Omega} G(\mathbf{x}, \mathbf{x}'; t, t') f(\mathbf{x}', t') \, d\mathbf{x}' dt'. \quad (2.82)$$

The solution derived at Equation 2.82 can be applied to Equation 2.33 obtaining,

$$\phi(\mathbf{x}, t) = \int_0^T \int_{\Omega} G(\mathbf{x}, \mathbf{x}'; t, t') S(\mathbf{x}', t') \, d\mathbf{x}' dt'. \quad (2.83)$$

where the Green function must fulfill the prescribed boundary conditions.

Now, let us subtract the Diffusion Approximation equation for an arbitrary absorption distribution with the Diffusion Approximation equation for a slightly perturbed absorption distribution,

$$\frac{1}{c} \frac{\partial \delta\phi}{\partial t} - \nabla \cdot (D\nabla\delta\phi) + \mu_a\phi - \mu_{a,0}\phi^0 = 0, \quad (2.84)$$

where  $\mu_{a,0}$  and  $\mu_a$  are the arbitrary and slightly changed absorption distributions ( $\delta\mu_a = \mu_a - \mu_{a,0}$ ),  $\phi^0$  and  $\phi$  are their respective fluence rates and  $\delta\phi = \phi - \phi^0$ . Acknowledging that

$$\mu_a\phi - \mu_{a,0}\phi^0 = \mu_a\phi - \mu_{a,0}(\phi - \delta\phi) = \delta\mu_a\phi + \mu_{a,0}\delta\phi, \quad (2.85)$$

Equation 2.84 transforms to

$$\frac{1}{c} \frac{\partial \delta\phi}{\partial t} - \nabla \cdot (D\nabla \delta\phi) + \mu_{a,0} \delta\phi = -\delta\mu_a\phi. \quad (2.86)$$

The solution of Equation 2.86 is:

$$\delta\phi(\mathbf{x}, \mathbf{x}_s; t) = - \int_0^T \int_{\Omega} G^0(\mathbf{x}, \mathbf{x}'; t, t') \delta\mu_a(\mathbf{x}') \phi(\mathbf{x}', \mathbf{x}_s; t') \, d\mathbf{x}' dt', \quad (2.87)$$

where  $\mathbf{x}_s$  is the source position and  $G^0$  is the Green's function for Diffusion Approximation with  $\mu_{a,0}$ . At detector position  $\mathbf{x}_d$  the difference in the fluence rate is,

$$\delta\phi(\mathbf{x}_d, \mathbf{x}_s; t) = - \int_0^T \int_{\Omega} G^0(\mathbf{x}_d, \mathbf{x}'; t, t') \delta\mu_a(\mathbf{x}') \phi(\mathbf{x}', \mathbf{x}_s; t') \, d\mathbf{x}' dt'. \quad (2.88)$$

The last equation may be rewritten as

$$\phi(\mathbf{x}_d, \mathbf{x}_s; t) = \phi^0(\mathbf{x}_d, \mathbf{x}_s; t) - \int_0^T \int_{\Omega} G^0(\mathbf{x}_d, \mathbf{x}'; t, t') \delta\mu_a(\mathbf{x}') \phi(\mathbf{x}', \mathbf{x}_s; t') \, d\mathbf{x}' dt', \quad (2.89)$$

where last term from right side integral can be expanded:

$$\begin{aligned} \phi(\mathbf{x}_d, \mathbf{x}_s; t) &= \phi^0(\mathbf{x}_d, \mathbf{x}_s; t) - \int_0^T \int_{\Omega} G^0(\mathbf{x}_d, \mathbf{x}'; t, t') \delta\mu_a(\mathbf{x}') \phi^0(\mathbf{x}', \mathbf{x}_s; t') \, d\mathbf{x}' dt' \\ &\quad + \int_0^T \int_{\Omega} \int_0^T \int_{\Omega} G^0(\mathbf{x}_d, \mathbf{x}'; t, t') \delta\mu_a(\mathbf{x}') G^0(\mathbf{x}', \mathbf{x}''; t', t'') \delta\mu_a(\mathbf{x}'') \\ &\quad \phi(\mathbf{x}'', \mathbf{x}_s; t'') \, d\mathbf{x}'' dt'' d\mathbf{x}' dt'. \end{aligned} \quad (2.90)$$

This expansion can be performed infinitely to get an infinite series which is known as *Born series*. In the following derivations two series truncation will be developed: Born and Rytov approximations.

### Born approximation

By truncating  $\delta\mu_a$  second and higher order terms of the Born series from Equation 2.90, the *Born approximation* is obtained:

$$\begin{aligned} \delta\phi(\mathbf{x}, \mathbf{x}_s; t) &\approx - \int_0^T \int_{\Omega} G^0(\mathbf{x}, \mathbf{x}'; t, t') \delta\mu_a(\mathbf{x}') \phi^0(\mathbf{x}', \mathbf{x}_s; t') \, d\mathbf{x}' dt' \\ &= - \int_0^T \int_{\Omega} G^0(\mathbf{x}, \mathbf{x}'; t - t') \delta\mu_a(\mathbf{x}') \phi_s^0(\mathbf{x}', t') \, d\mathbf{x}' dt' \\ &= - \int_{\Omega} [G^0(\mathbf{x}, \mathbf{x}') * \phi_s^0(\mathbf{x}')](t) \delta\mu_a(\mathbf{x}') \, d\mathbf{x}', \end{aligned} \quad (2.91)$$

where  $\phi^0(\mathbf{x}', \mathbf{x}_s; t') = \phi_s^0(\mathbf{x}', t')$ .

Born approximation can be interpreted as the first order approximation of the Born series. At detector position the Born approximation is,

$$\begin{aligned}\delta\phi_{sd} &\approx - \int_0^T \int_{\Omega} G^0(\mathbf{x}_d, \mathbf{x}'; t - t') \delta\mu_a(\mathbf{x}') \phi_s^0(\mathbf{x}', t') d\mathbf{x}' dt' \\ &= - \int_{\Omega} [G_d^0(\mathbf{x}') * \phi_s^0(\mathbf{x}')](t) \delta\mu_a(\mathbf{x}') d\mathbf{x}',\end{aligned}\quad (2.92)$$

where  $\delta\phi_{sd} = \delta\phi(\mathbf{x}_d, \mathbf{x}_s; t)$  and  $G_d^0(\mathbf{x}') = G^0(\mathbf{x}_d, \mathbf{x}') = G^0(\mathbf{x}', \mathbf{x}_d)$  because of symmetry of Green's function and  $*$  is the convolution operator.

Born approximation can also be extended for diffusion coefficient heterogeneities (see Appendix E.1),

$$\delta\phi_{sd} \approx - \int_{\Omega} (\phi_s^0 * G_d^0) \delta\mu_a d\mathbf{x} - \int_{\Omega} (\nabla\phi_s^0 * \nabla G_d^0) \delta D d\mathbf{x}.\quad (2.93)$$

This equation is the cornerstone of DOT; it tells us how much absorption and diffusion coefficient distribution should be changed,  $\delta\mu_a(\mathbf{x})$  and  $\delta D(\mathbf{x})$ , based on the detector signal difference between unknown and simulated medium,  $\delta\phi_{sd}$ . This approximation can be discretized and solved for  $\delta\mu_a$  and  $D$ ,

$$\begin{bmatrix} \delta\phi_{11} \\ \delta\phi_{12} \\ \vdots \\ \delta\phi_{N_s N_d} \end{bmatrix} \approx \begin{bmatrix} \ddots & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots \end{bmatrix} \begin{bmatrix} (\delta\mu_a)_1 \\ (\delta\mu_a)_2 \\ \vdots \\ (\delta\mu_a)_M \end{bmatrix} + \begin{bmatrix} \ddots & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots \end{bmatrix} \begin{bmatrix} (\delta D)_1 \\ (\delta D)_2 \\ \vdots \\ (\delta D)_M \end{bmatrix},\quad (2.94)$$

which can be combined into a single linear system:

$$\begin{bmatrix} \delta\phi_{11} \\ \delta\phi_{12} \\ \vdots \\ \delta\phi_{N_s N_d} \end{bmatrix} \approx \begin{bmatrix} \ddots & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \ddots \end{bmatrix} \begin{bmatrix} (\delta\mu_a)_1 \\ (\delta\mu_a)_2 \\ \vdots \\ (\delta\mu_a)_M \\ (\delta D)_1 \\ (\delta D)_2 \\ \vdots \\ (\delta D)_M \end{bmatrix}.\quad (2.95)$$

### Rytov approximation

For any single variable function  $f$  it holds,

$$\ln(f(x+h)) - \ln(f(x)) = \ln\left(\frac{f(x+h)}{f(x)}\right) = \ln\left(\frac{f(x) + \delta f}{f(x)}\right) = \ln\left(1 + \frac{\delta f}{f(x)}\right).\quad (2.96)$$

The Taylor expansion of the exponential function around  $x = 0$  is

$$e^x = 1 + x + \frac{x^2}{2!} + \dots,\quad (2.97)$$

and from its natural logarithm it follows that

$$x = \ln \left( 1 + x + \frac{x^2}{2!} + \dots \right) = \ln(1 + x) + \mathcal{O}(x^2). \quad (2.98)$$

If  $x$  is very small, second order terms can be neglected (see Figure 2.16),

$$x \approx \ln(1 + x). \quad (2.99)$$

Therefore, equation

$$\ln(f(x + h)) - \ln(f(x)) = \ln \left( 1 + \frac{\delta f}{f(x)} \right), \quad (2.100)$$

can be approximated by

$$\ln(f(x + h)) - \ln(f(x)) \approx \frac{\delta f}{f(x)}, \quad (2.101)$$

assuming  $\delta f/f$  is small.

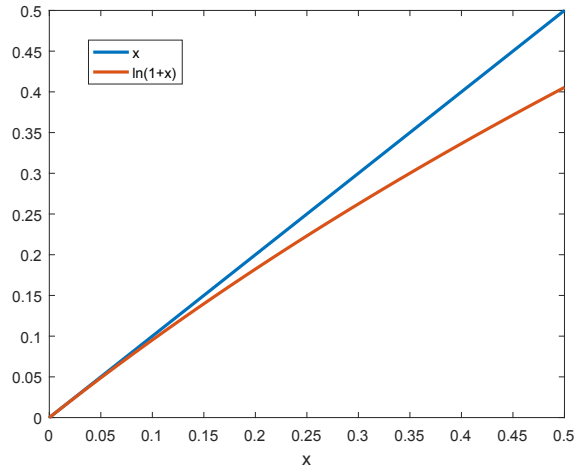


Figure 2.16: Validity of  $x \approx \ln(1 + x)$  approximation.

Using the approximation above, the Born approximation can be converted to the well-known Rytov approximation,

$$\begin{aligned} \frac{\delta \phi_{sd}}{\phi_{sd}^0} &= -\frac{1}{\phi_{sd}^0} \int_{\Omega} [G_d^0(\mathbf{x}') * \phi_s^0(\mathbf{x}')(t)] \delta \mu_a(\mathbf{x}') d\mathbf{x}' \\ \Rightarrow \ln \left( \frac{\phi_{sd}}{\phi_{sd}^0} \right) &\approx -\frac{1}{\phi_{sd}^0} \int_{\Omega} [G_d^0(\mathbf{x}') * \phi_s^0(\mathbf{x}')(t)] \delta \mu_a(\mathbf{x}') d\mathbf{x}', \end{aligned} \quad (2.102)$$

Rytov approximation can be interpreted as an approach to model the second and higher order terms of the Born series expansion using the natural logarithm.



### Pseudocode of the reconstruction algorithm

The linearization of the Born approximation described at Equation 2.92 is,

$$\delta\phi_{sd} = \mathbf{J}\delta\mu_a, \quad (2.103)$$

where matrix  $\mathbf{J}$  is known as the jacobian or sensitivity matrix and is a discretization of  $-(G_d * \phi_s)$  multiplied by voxel volume factors. An iterative reconstruction algorithm can be developed as follows:

$N_{\max}$  = maximum number of iterations

Set initial guess of absorption coefficient  $\mu_a^0(\mathbf{x})$  (usually uniform)

For loop ( $k = 0 : N_{\max}$ ):

1. Solve forward problem with  $\mu_a^k(\mathbf{x})$  from iteration  $k$ -th.
2. Set  $\delta\phi_{sd}$ : signal difference between unknown and simulated medium.
3. Compute the jacobian matrix  $\mathbf{J}$ .
4. Solve linear system  $\delta\phi_{sd} = \mathbf{J}\delta\mu_a$  by applying some regularization.
5.  $\mu_a^{k+1}(\mathbf{x}) = \mu_a^k(\mathbf{x}) + \delta\mu_a^k(\mathbf{x})$

If  $\|\mu_a^{k+1} - \mu_a^k\|_2 / \|\mu_a^{k+1}\|_2 < \varepsilon$  then convergence has been reached and the algorithm is stopped.

As can be seen, Born approximation is used iteratively until convergence is reached. That is, at first iteration absorption is estimated by  $\mu_a^1(\mathbf{x}) = \mu_a^0(\mathbf{x}) + \delta\mu_a^0(\mathbf{x})$ . After, if convergence is not reached then  $\delta\mu_a^1$  is computed and absorption updated again. This procedure is processed until the absorption relative update is smaller than threshold  $\varepsilon$ , or the maximum number of iterations has been completed.

### 2.5.2. Reconstruction using continuous-wave systems

In the continuous-wave case, the measurements are constant over time. The Born approximation simplifies to

$$\phi_{sd} \approx \phi_{sd}^0 - \int_{\Omega} \phi_s^0(\mathbf{x}) G_d^0(\mathbf{x}) \delta\mu_a(\mathbf{x}) d\mathbf{x}. \quad (2.104)$$

If the approximation is discretized in voxels, the explicit form of the linear system is,

$$\begin{bmatrix} \delta\phi_{11} \\ \delta\phi_{12} \\ \vdots \\ \delta\phi_{SD} \end{bmatrix} = - \begin{bmatrix} V_1\phi_1^0(x_1)G_1^0(x_1) & V_2\phi_1^0(x_2)G_1^0(x_2) & \cdots & V_M\phi_1^0(x_M)G_1^0(x_M) \\ V_1\phi_1^0(x_1)G_2^0(x_1) & V_2\phi_1^0(x_2)G_2^0(x_2) & \cdots & V_M\phi_1^0(x_M)G_2^0(x_M) \\ \vdots & \vdots & \cdots & \vdots \\ V_1\phi_S^0(x_1)G_D^0(x_1) & V_2\phi_S^0(x_2)G_D^0(x_2) & \cdots & V_M\phi_S^0(x_M)G_D^0(x_M) \end{bmatrix} \begin{bmatrix} (\delta\mu_a)_1 \\ (\delta\mu_a)_2 \\ \vdots \\ (\delta\mu_a)_M \end{bmatrix} \quad (2.105)$$

where  $\mathbf{J} \in \mathbb{R}^{(S \cdot D) \times M}$ ,  $\delta\mu_a \in \mathbb{R}^M$  and  $\delta\phi_{sd} \in \mathbb{R}^{(S \cdot D)}$  being  $S \cdot D$  the number of source-detectors pairs and  $M$  the number of voxels in the domain. The term  $V_i$  corresponds to the area (or volume in 3D) of voxel  $i$ -th.

The jacobian matrix  $\mathbf{J}$  is usually ill-posed and not squared. By multiplying both parts of the equation by the transpose  $\mathbf{J}^T$ ,

$$\mathbf{J}^T \mathbf{J} \delta \mu_a = \mathbf{J}^T \delta \phi_{sd}, \quad (2.106)$$

a square matrix is obtained. The matrix is still ill-posed and regularization has to be applied. A commonly used and easy to implement method is the Tikhonov regularization [102], applied by the normal equation,

$$\mathbf{J}^T \mathbf{J} (\delta \mu_a + \lambda \mathbf{I}) = \mathbf{J}^T \delta \phi_{sd}, \quad (2.107)$$

where  $\lambda$  is the regularization parameter and  $\mathbf{I}$  is the identity matrix. Note that the normal equation is equivalent to the equation solved by the Levenberg-Marquardt algorithm. The jacobian matrix can be computed using the adjoint form [5] which needs to solve  $N_s + N_d$  different forward problems. For noisy measurements, the linear system has to be multiplied by a weight matrix, in order not to overfit the result. The jacobian could be redefined as  $\mathbf{J} = \mathbf{W} \mathbf{J}$  where  $\mathbf{W}$  is diagonal matrix with components  $W_{ii} = 1/\sqrt{\sigma_i^2}$  where  $\sigma_i$  is the standard deviation of measurement  $i$ -th.

### 2.5.3. Reconstruction using frequency-resolved systems

Frequency-resolved systems measure the magnitude and phase at different frequencies. The linear system to solve is:

$$\begin{bmatrix} \text{Mag} [\delta \phi_{sd}(\omega_1)] \\ \text{Pha} [\delta \phi_{sd}(\omega_1)] \\ \text{Mag} [\delta \phi_{sd}(\omega_2)] \\ \text{Pha} [\delta \phi_{sd}(\omega_2)] \\ \vdots \\ \text{Mag} [\delta \phi_{sd}(\omega_N)] \\ \text{Pha} [\delta \phi_{sd}(\omega_N)] \end{bmatrix} = \begin{bmatrix} \text{Mag} [J_{sd}(\omega_1, x_1)] & \cdots & \text{Mag} [J_{sd}(\omega_1, x_M)] \\ \text{Pha} [J_{sd}(\omega_1, x_1)] & \cdots & \text{Pha} [J_{sd}(\omega_1, x_M)] \\ \text{Mag} [J_{sd}(\omega_2, x_1)] & \cdots & \text{Mag} [J_{sd}(\omega_2, x_M)] \\ \text{Pha} [J_{sd}(\omega_2, x_1)] & \cdots & \text{Pha} [J_{sd}(\omega_2, x_M)] \\ \vdots & & \vdots \\ \text{Mag} [J_{sd}(\omega_N, x_1)] & \cdots & \text{Mag} [J_{sd}(\omega_N, x_M)] \\ \text{Pha} [J_{sd}(\omega_N, x_1)] & \cdots & \text{Pha} [J_{sd}(\omega_N, x_M)] \end{bmatrix} \begin{bmatrix} (\delta \mu_a)_1 \\ \vdots \\ (\delta \mu_a)_M \end{bmatrix}, \quad (2.108)$$

where  $J_{sd}(\cdot, \cdot)$  and  $\delta \phi_{sd}(\cdot) \in \mathbb{C}^{S \cdot D}$  and  $\text{Mag}[\cdot]$  and  $\text{Pha}[\cdot]$  are the magnitude and phase at angular frequency  $\omega$ . In this case the system matrix size is  $(2N \cdot S \cdot D) \times M$  where  $N$  is the number of angular frequencies used. (Note that continuous-wave is the specific  $\omega = 0$  case).

### 2.5.4. Reconstruction using time-resolved systems

Equation 2.92 can be directly discretized into a time-resolved linear system,

$$\begin{bmatrix} \delta \phi_{sd}(t_1) \\ \delta \phi_{sd}(t_2) \\ \vdots \\ \delta \phi_{sd}(t_N) \end{bmatrix} = \begin{bmatrix} J_{sd}(t_1, x_1) & J_{sd}(t_1, x_2) & \cdots & J_{sd}(t_1, x_M) \\ J_{sd}(t_2, x_1) & J_{sd}(t_2, x_2) & \cdots & J_{sd}(t_2, x_M) \\ \vdots & \vdots & \vdots & \vdots \\ J_{sd}(t_N, x_1) & J_{sd}(t_N, x_2) & \cdots & J_{sd}(t_N, x_M) \end{bmatrix} \begin{bmatrix} \delta \mu_{a,1} \\ \delta \mu_{a,2} \\ \vdots \\ \delta \mu_{a,M} \end{bmatrix}, \quad (2.109)$$

where  $J_{sd}(\cdot, \cdot)$  and  $\delta \phi_{sd}(\cdot) \in \mathbb{R}^{S \cdot D}$ .

In this case the system matrix size is  $(N \cdot S \cdot D) \times M$ , where  $N$  is the number of time bins considered. Posing the reconstruction directly in terms of time has two main drawbacks:

(1) the number of time bins is quite high (usually around one thousand), which makes the system very large and (2) it is necessary to compute the TPSF for each source and detector which is highly time-consuming. Due to these drawbacks usually datatypes (also known as temporal filters or windows) [193] are used instead. These datatypes are just windows that are applied to the DTOF in order to compress the information and to promote fast computing. In the literature mainly three kind of datatypes are used: standard moments, Mellin–Laplace windows and Fourier transform. The former two will be addressed in the following subsections, the latter was introduced in the previous subsection.

### Standard moments

The  $n$ -th moment of a function  $f(t)$  is defined as

$$[f(t)]^n = \int_{-\infty}^{\infty} f(t)t^n dt, \quad (2.110)$$

which are known as *standard moments*. Since the  $n$ -th moment of a convolution of two functions is

$$[f(t) * g(t)]^n = \sum_{k=0}^n \binom{n}{k} [f(t)]^k [g(t)]^{n-k}, \quad (2.111)$$

where  $\binom{n}{k} = n!/(k!(n-k)!)$  the  $n$ -th moment of Born approximation is

$$[\delta\phi_{sd}(t)]^n \approx - \int_{\Omega} \sum_{k=0}^n \binom{n}{k} [\phi_s(\mathbf{x}, t)]^k [G_d(\mathbf{x}, t)]^{n-k} \delta\mu_a(\mathbf{x}) d\mathbf{x}. \quad (2.112)$$

Standard moments can be computed directly without computing explicitly the TPSF. The following derivation is based on the calculations developed at [7]. The  $n$ -th derivative of a Fourier transform is

$$\frac{\partial^n}{\partial \omega^n} \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt = (-i)^n \int_{-\infty}^{\infty} t^n f(t)e^{-i\omega t} dt. \quad (2.113)$$

If only the frequency  $\omega = 0$  is taken and using the identity  $i^n(-i)^n = 1$ , the previous equation transforms to

$$i^n \frac{\partial^n}{\partial \omega^n} \hat{f}(\omega) \Big|_{\omega=0} = [f(t)]^n, \quad (2.114)$$

where  $[f(t)]^n$  is the  $n$ -th moment of function  $f(t)$  and  $\hat{f}$  is its Fourier transform. The Fourier transform of the time-dependant linear system (Equation 2.49) is

$$[i\omega\mathbf{H} + \mathbf{M} + \mathbf{K} + \mathbf{N}]\hat{\Psi} = \hat{\mathbf{s}}, \quad (2.115)$$

where  $\hat{\Psi}$  represents the Fourier transform of function  $\Psi$ . Imposing a pulse source,

$$[i\omega\mathbf{H} + \mathbf{M} + \mathbf{K} + \mathbf{N}]\hat{G} = \delta(\mathbf{x} - \mathbf{x}_s)e^{-i\omega t_s} = \delta(\mathbf{x} - \mathbf{x}_s), \quad (2.116)$$

is obtained. Its  $n$ -th derivative with respect to  $\omega$  is,

$$[\mathbf{M} + \mathbf{K} + \mathbf{N}]\frac{\partial^n \hat{G}}{\partial \omega^n} + in\mathbf{H}\frac{\partial^{n-1} \hat{G}}{\partial \omega^{n-1}} + i\omega\mathbf{H}\frac{\partial^n \hat{G}}{\partial \omega^n} = 0. \quad (2.117)$$

Moreover, if only the  $\omega = 0$  frequency is taken, both sides are multiplied by  $i^n$  and identity from Equation 2.114 is used the following linear system is obtained,

$$(\mathbf{M} + \mathbf{K} + \mathbf{N})[G(\mathbf{x}, t)]^n = n\mathbf{H}[G(\mathbf{x}, t)]^{n-1}. \quad (2.118)$$

Therefore, the linear system above computes directly the  $n$ -th standard moment of Green's function by using the  $(n - 1)$ -th moment. This implies that standard moments can be computed iteratively without computing explicitly the TPSF.

### Mellin–Laplace moments

The Mellin–Laplace (ML) transform of a function  $f(t)$  is defined as

$$[f(t)]^{(n,p)} = \frac{p^n}{n!} \int_0^\infty f(t)t^n e^{-pt} dt, \quad n \in \mathbb{N}, \quad p \in \mathbb{R}. \quad (2.119)$$

The ML moment of a convolution is given by:

$$[f(t) * g(t)]^{(n,p)} = \sum_{k=0}^n [f(t)]^{(k,p)} [g(t)]^{(n-k,p)}, \quad (2.120)$$

and for a given moment  $n$  and a value  $p$  the Born approximation transforms to

$$[\delta\phi_{sd}]^{(n,p)} = - \int_{\Omega} \sum_{k=0}^n [\phi_s(\mathbf{x}, t)]^{(k,p)} [G_d(\mathbf{x}, t)]^{(n-k,p)} \delta\mu_a(\mathbf{x}) d\mathbf{x}. \quad (2.121)$$

Mellin–Laplace moments can also be computed using FEM matrices without computing explicitly the TPSF signal. Similar procedure as with standard moments can be used. The  $n$ -th derivative of a Laplace transform is,

$$\frac{d^n}{d\rho^n} \int_0^\infty f(t)e^{-\rho t} dt = (-1)^n \int_0^\infty t^n f(t)e^{-\rho t} dt, \quad (2.122)$$

and after some algebra, the Mellin–Laplace transform is equal to,

$$\frac{\rho^n}{n!(-1)^n} \frac{d^n}{d\rho^n} \int_0^\infty f(t)e^{-\rho t} dt = \frac{\rho^n}{n!} \int_0^\infty t^n f(t)e^{-\rho t} dt = f^{(p,n)}. \quad (2.123)$$

The Laplace transform of the time-dependant linear systems is,

$$[\rho\mathbf{H} + \mathbf{M} + \mathbf{K} + \mathbf{N}]\tilde{\Psi} = \tilde{\mathbf{s}}. \quad (2.124)$$

where  $\tilde{\Psi}$  is the Laplace transform of function  $\Psi$ .

If the  $n$ -th derivative is applied to Equation (2.124) and Dirac delta source is assumed, then

$$[\mathbf{M} + \mathbf{K} + \mathbf{N}] \frac{d^n \tilde{G}}{d\rho^n} + n\mathbf{H} \frac{d^{n-1} \tilde{G}}{d\rho^{n-1}} + \rho\mathbf{H} \frac{d^n \tilde{G}}{d\rho^n} = 0. \quad (2.125)$$

Multiplying by  $\rho^n / (n!(-1)^n)$  and using the identity from Equation 2.123, the following linear system is derived

$$[\mathbf{M} + \mathbf{K} + \mathbf{N} + \rho\mathbf{H}]G^{(\rho,n)} = \rho\mathbf{H}G^{(\rho,n-1)}, \quad (2.126)$$

where  $G^{(\rho,0)}$  is the Laplace transform of  $G$ , that is, the solution to Equation (2.124) for a pulse source. Therefore,  $G^{(\rho,n)}$  is given by solving the linear system with  $G^{(\rho,n-1)}$ .

### 2.5.5. Measurements into the reconstruction algorithms

The measurements need to be introduced into the Born approximation described at Equation 2.92. Measurements can be defined as,

$$M_{sd} = G_{sd} * S(t) * D(t) + \eta(t) \quad (2.127)$$

where  $G_{sd}(t)$  is the theoretical signal (free of noise and instrumental factors; the source is assumed to be a Dirac's delta distribution),  $S(t)$  accounts for source instrumental factors (time delays, power and broadness),  $D(t)$  accounts for detector instrumental factors (time delays and detection efficiency) and  $\eta(t)$  is the noise term. For simplicity, the noise influence will be ignored from now (in Chapter 3 is given an extensive analysis of noise, its influence and treatment for tomography problems).

The experimental factors can be included into the Born approximation as follows:

$$\delta G_{sd} * S(t) * D(t) \approx -S(t) * D(t) * \int_0^T \int_{\Omega} G^0(\mathbf{x}_d, \mathbf{x}'; t - t') \delta\mu_a(\mathbf{x}') G_s^0(\mathbf{x}', t') d\mathbf{x}' dt', \quad (2.128)$$

where  $\phi$  was made equal to  $G * S(t) * D(t)$  since it was assumed that the differences between the Dirac's delta and light pulse were included inside  $S(t)$ .

From Equation 2.128, there exist two different approaches to perform the reconstruction: (1) to use the cross approach described in [101] where a known medium is used as reference and (2) to deconvolve instrumental factors from the signals. Here, I will describe both techniques, where in both cases it can be assumed that the IRF is unknown if a measurement is done in a reference medium. In Chapter 3, I will explain a different reconstruction technique which deconvolves the instrumental factors directly.

To apply the cross-Born approximation two types of measurements will be considered:  $M_{sd}^A$  where  $A$  indicates the reference medium which optical properties are known and  $M_{sd}^B$  where  $B$  indicates the target unknown medium (see Figure 2.17).

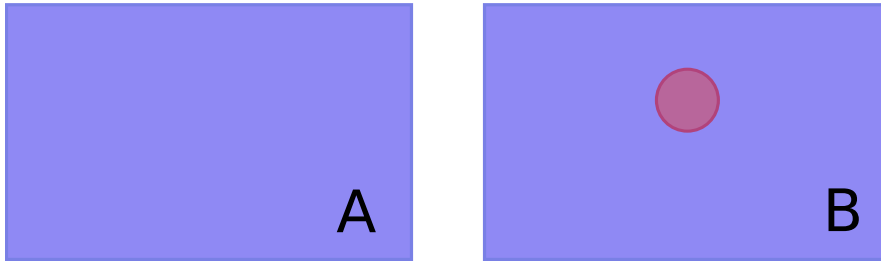


Figure 2.17: (Left) Reference medium which optical properties are known; usually an homogeneous medium is used. (Right) Target medium whose optical properties are unknown; the red circle represents an absorption heterogeneity.

Using some algebra,

$$\begin{aligned} M_{sd}^B * G_{sd}^A - M_{sd}^A * G_{sd}^{B(k)} &= G_{sd}^B * \underbrace{G_{sd}^A * S * D}_{M_{sd}^A} - G_{sd}^{B(k)} * \underbrace{G_{sd}^A * S * D}_{M_{sd}^A} \\ &= \underbrace{(G_{sd}^B - G_{sd}^{B(k)})}_{\delta G_{sd}^B} * M_{sd}^A = \delta G_{sd}^B * M_{sd}^A, \end{aligned} \quad (2.129)$$

where  $G_{sd}^{B(k)}$  denotes the simulated signal for the absorption recovered at iteration  $k$ -th assuming a Dirac's delta distribution as source.

So, the relationship

$$M_{sd}^B * G_{sd}^A - M_{sd}^A * G_{sd}^{B(k)} = \delta G_{sd}^B * M_{sd}^A, \quad (2.130)$$

can be introduced at Born approximation (Equation 2.128) obtaining

$$M_{sd}^B * G_{sd}^A - M_{sd}^A * G_{sd}^{B(k)} = -M_{sd}^A * \int_{\Omega} \left[ G_s^{B(k)}(\mathbf{x}, t) * G_d^{B(k)}(\mathbf{x}, t) \right] \delta\mu_a(\mathbf{x}) \, d\mathbf{x}, \quad (2.131)$$

which is known as the *cross-Born approximation*. Since Born approximation and measurements have been related in a single equation, now reconstruction can be performed from experimental data. Nevertheless, measurements and simulations were convolved which implies that some information will be lost as proven in Chapter 3. Noise factor can be overcome by using the generalized least squares method (see Appendix D) in combination with the covariance matrix.

A second approach, that does not convolve the data, is to first deconvolve the instrumental factors  $S * D$  from the measurements and introduce them into the plain Born approximation. The technique explained below assumes that  $S * D$  is unknown. Mellin–Laplace transform can be applied to the reference measurement

$$M_{sd}^A = G_{sd}^A * S * D \xrightarrow{\text{M-L}} [M_{sd}^A]^{(n,p)} = \sum_{i=0}^n [G_{sd}^A]^{(i,p)} [S * D]^{(n-i,p)} \quad (2.132)$$

and rewritten in matrix form by using a Toeplitz matrix

$$\begin{bmatrix} [M_{sd}^A]^{(0,p)} \\ [M_{sd}^A]^{(1,p)} \\ [M_{sd}^A]^{(2,p)} \\ \vdots \\ [M_{sd}^A]^{(n,p)} \end{bmatrix} = \begin{bmatrix} [G_{sd}^A]^{(0,p)} & 0 & 0 & \cdots & 0 \\ [G_{sd}^A]^{(1,p)} & [G_{sd}^A]^{(0,p)} & 0 & \cdots & 0 \\ [G_{sd}^A]^{(2,p)} & [G_{sd}^A]^{(1,p)} & [G_{sd}^A]^{(0,p)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ [G_{sd}^A]^{(n,p)} & [G_{sd}^A]^{(n-1,p)} & [G_{sd}^A]^{(n-2,p)} & \cdots & [G_{sd}^A]^{(0,p)} \end{bmatrix} \begin{bmatrix} [S * D]^{(0,p)} \\ [S * D]^{(1,p)} \\ [S * D]^{(2,p)} \\ \vdots \\ [S * D]^{(n,p)} \end{bmatrix}. \quad (2.133)$$

Mellin–Laplace moments of  $S * D$  can be computed by solving this linear system. Measurements for medium  $B$  can also be represented by a Toeplitz matrix,

$$\begin{bmatrix} [M_{sd}^B]^{(0,p)} \\ [M_{sd}^B]^{(1,p)} \\ [M_{sd}^B]^{(2,p)} \\ \vdots \\ [M_{sd}^B]^{(n,p)} \end{bmatrix} = \begin{bmatrix} [G_{sd}^B]^{(0,p)} & 0 & 0 & \cdots & 0 \\ [G_{sd}^B]^{(1,p)} & [G_{sd}^B]^{(0,p)} & 0 & \cdots & 0 \\ [G_{sd}^B]^{(2,p)} & [G_{sd}^B]^{(1,p)} & [G_{sd}^B]^{(0,p)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ [G_{sd}^B]^{(n,p)} & [G_{sd}^B]^{(n-1,p)} & [G_{sd}^B]^{(n-2,p)} & \cdots & [G_{sd}^B]^{(0,p)} \end{bmatrix} \begin{bmatrix} [S * D]^{(0,p)} \\ [S * D]^{(1,p)} \\ [S * D]^{(2,p)} \\ \vdots \\ [S * D]^{(n,p)} \end{bmatrix}, \quad (2.134)$$

where, in this case, the Toeplitz matrix is unknown but it can be computed by solving each equation from the top to the bottom sequentially.

That is,

$$[G_{sd}^B]^0 = \frac{[M_{sd}^B]^0}{[S * D]^0} \quad (2.135)$$

$$[G_{sd}^B]^1 = \frac{[M_{sd}^B]^1 - [G_{sd}^B]^0 [S * D]^1}{[S * D]^0} \quad (2.136)$$

which can be generalized to

$$[G_{sd}^B]^n = \frac{[M_{sd}^B]^n - \sum_{i=1}^n [G_{sd}^B]^{n-i} [S * D]^i}{[S * D]^0}. \quad (2.137)$$

As can be seen, the main drawback of this method is that noise is propagated along each order (see Equation 2.137).

## 2.6. Clinical perspectives: applications to neuromonitoring

After presenting the fundamentals and basics of diffuse optics, it is important to consider the potential clinical applications. Frans Jöbsis in 1977 was the first to measure oxy and deoxyhemoglobin changes in a mammal [112] by using near-infrared spectroscopy (NIRS). After, the scientific community has studied these hemoglobin changes at different human organs in order to diagnose or monitorize several diseases such as thrombosis [125], different types of cancer [121], flap monitoring [37, 58, 173] or brain imaging [228]. Moreover, NIRS have also been used to monitor patients during surgeries, for example cerebral tissue oxygen saturation was measured in cardiac operations [35, 78, 155]. Nevertheless, there is still an ongoing debate about NIRS benefits during cardiac surgery [20, 199, 157] and, therefore more research have to be done in the clinics.

In the next subsections, I will introduce three neuromonitoring applications that have a large potential to be used systematically in the clinics of the future. In fact, some of them have already been tested in some clinical situations with promising results.

### 2.6.1. Traumatic brain injuries

NIRS has a big potential in the context of traumatic brain injury (TBI). The clinicians identify two cases where NIRS could be used successfully [52]: 1) as a bedside continuous monitoring tool for the detection of secondary brain injuries (primary brain injuries are generated by the concussion or accident and secondary are a produced by the increase of intracranial pressure due to the primary hematoma) and 2) as a tool for initial diagnostic in primary care settings or emergency situations. It must be noted that the idea is not to replace the well-established CT or MRI imaging techniques, since they are still the best choice at secondary and tertiary care settings. The plan is to use NIRS technology for doing a fast pre-evaluation just after the injury has occurred (e.g. sport related concussions [75]) or to do a continuous monitoring where CT and MRI are not affordable or appropriate. In this context, the detection of acute intracranial hematomas right after the injury has occurred (usually at out-of-hospital situations or in non-specialized medical centers) would be very helpful for doctors in order to decide if a patient should be send to a secondary care setting or not [52, 198].

Several attempts have been done to detect hematomas by using NIRS [186, 114, 187], all of them showed high sensitivity and specificity levels. Moreover, Infrascanner device has been commercialized recently for screening in-situ the presence of brain hematomas [124]; Infrascanner was used by the German army in Afghanistan [27]. However, some of the expectations have been decreased due to several practical limitations [235, 198], such as contamination of signal by extracranial tissue (scalp, skull and cerebrospinal fluid), the difficulty of obtaining absolute values or inaccuracies when hematomas exist in both hemispheres. Therefore, there

are still many issues that need to be addressed before fully applying NIRS into the clinics for TBI monitoring and assessment.

### 2.6.2. Newborn hypoxia

Hypoxic ischemic encephalopathies, such as cerebral palsy [225], are one of the leading causes of death and long-term neurological deficits in neonates. Although MRI or CT are gold standard techniques in adult patients these tools are difficult or even risky to use in few-days old babies [11, 16]. For these reasons, there is a need of a technology that could allow to perform bedside, continuous, non-invasive and safe brain monitoring in newborn patients. Among the technologies that fulfill these properties optical based systems are one of the most promising techniques.

NIRS technology was one of the first methods to be proposed [240, 209] as an alternative technique for neonates since their skull and scalp are less thick than in adults and therefore measurements are less noisy. Moreover, some studies have validated NIRS in neonates by reporting a good correlation with MRI [238]. NIRS have also been proposed as an alternative for the monitoring of neonatal brain injuries [13, 149]. Nowadays, although there are many commercial devices that measure the trends of cerebral oxygen saturation the absolute values are still a challenge [59] and are the major obstacle to fully implement NIRS into the newborn health care system. Therefore, in this field there is still a lot of space for improvement that could be exploited by, for example, finding some protocols or standards to homogenize the diverse results in terms of absolute oxygen saturation values.

### 2.6.3. Functional imaging

NIRS can also be used to analyze brain activity. When a part of the brain is activated the cerebral blood flow in that region will increase due to neurovascular coupling. The quantity of oxygenated blood that the flow carries is greater than the activated neurons need. Therefore, not all the oxygenated blood will be consumed and the local concentration of oxygenated blood will increase in overall. At the same time the blood flow increases in the active region the deoxygenated blood will decrease slightly. That is, when neural activity takes place, the concentration of oxygenated blood will increase but the concentration of deoxygenated blood will decrease slightly. Usually, these changes of concentration do not occur immediately but with a couple of seconds of delay. Functional imaging methods that measure these hemoglobin changes are known as blood-oxygen-level dependent imaging techniques or BOLD imaging techniques.

There already exist many well-established brain function imaging methods such as functional MRI (fMRI), electro encephalography (EEG) or magneto encephalography (MEG). fMRI is a BOLD imaging technique, but it measures the magnetic susceptibility of oxy and deoxygenated blood instead of the absorption. However, EEG and MEG are not BOLD techniques since they measure the neural activity in the brain by recording the electrical or magnetic potentials of neurons. Although all these methods are well-established in research and clinics they still require the patient to be restrained since any movement could produce artifacts in the measurements [189]. Nevertheless, several studies used NIRS to analyze brain activation of walking subjects, e.g. in healthy population [213], elderly population [93] or subject that suffered a stroke [147]. Moreover, since NIRS systems are quite inexpensive, small and



can be used in environments with high radiation or strong magnetic fields its usefulness at neuroscientific research is large and well-rooted.

In the next chapter, I develop a generalized theory for temporal windows datatypes. Strong emphasis is put on analysing different windows and how they are influence by instrumental factors. I consider that the work presented in following chapter could be applied to the clinical neuromonitoring devices that will be commercialized in the future.

# Theoretical study of datatypes for time-resolved diffuse optical tomography

**D**IFFUSE optical tomography recovers the optical properties of an unknown diffusive medium by solving an ill-posed inverse problem. In time-resolved technology, reconstructions based on datatypes (e.g. temporal windows) are usually used because they are faster to compute than the fluence rate. Nevertheless, neither theoretical nor numerical studies assessing different datatypes have been clearly expressed.

In this chapter, I propose an overview and a new process to compute efficiently a long set of temporal windows in order to perform diffuse optical tomography. I carried out a theoretical comparison of these large set of temporal windows including also Fourier transform datatypes. I also did simulations in a reflectance geometry with a spherical inclusion at different depths. The results are presented in terms of inclusion localization and its absorption coefficient recovery. I show that (1) the new windows computed with the developed method perform better than the state-of-the-art windows for inclusions deeper than 2.5 cm, (2) in some cases these windows can be equivalent to frequency based reconstruction at GHz order and (3) deconvolving the instrumental response function of the measurements has a superior performance compared to the cross-Born inverse problem approach.

## 3.1. Introduction

With time-resolved diffuse optical technology the concentrations of several physiological chromophores such as hemoglobin, lipid or collagen can be measured in-vivo. Successful measurements have been done at different human body locations such as brain [235, 44], breast [215] or thyroid [132]. An interesting extension of this technology is to perform diffuse optical tomography [203, 9] by computing three-dimensional maps of oxy- and deoxy-hemoglobin; in this approach, photon propagation is modeled in a computer and results are compared with experimental measurements. Then, optical parameters in the model are updated by solving an ill-posed inverse problem until the difference between the model and

the data is negligible.

In order to have accurate results, it is important to have a realistic model for photon propagation in tissues. The most accurate approach is to use the integro-differential Radiative Transfer Equation (RTE) described at Subsection 2.2.2. Although RTE has some analytical solutions [130, 128, 129] these just hold for simple geometries and cannot be applied to more complex environments, such as an human head or breast models, without making strong assumptions. Apart from classical Monte-Carlo simulations [233, 68], new numerical methods have been proposed, some of which are the one-way RTE [84] or hybrid RTE [216]; nevertheless, they are still highly time-consuming for real-time applications.

Due to the previous reasons, usually the first-order time-dependant Diffusion Approximation, described at Equation 2.24 from Subsection 2.2.2, is used. This approximation, based on spherical harmonics [234], is valid assuming that (1) the reduced scattering coefficient is much greater than absorption ( $\mu'_s \gg \mu_a$ ) and that (2) detectors are sufficiently far from light sources ( $> 1/\mu'_s$ ). These assumptions hold in most practical cases and although higher order approximations have also been suggested [40, 136, 244] the extra computational cost usually does not compensate the small gain in accuracy.

The time-dependant Diffusion Approximation equation can be solved using Finite Element Method [10] but it is still a slow process since time has to be discretized in short steps in order to guarantee stability and convergence. For this reason, several authors have proposed to use temporal windows datatypes since they are faster to compute than the fluence rate itself [101, 7]. Nevertheless, as indicated in [5], the temporal windows datatypes are faster to compute than the fluence rate if they have the form  $w(t) = t^n e^{-pt}$ , where  $n \in \mathbb{N}$  is in the set of natural numbers and  $p \in \mathbb{C}$  is in the set of complex numbers. This formula incorporates the Fourier transform (when  $n = 0$  and  $p$  is an imaginary number), Laplace transform (when  $n = 0$  and  $p$  is a complex number), standard moments (when  $n$  is an integer and  $p = 0$ ) and Mellin–Laplace moments (when  $n$  is an integer and  $p$  is a real number).

Some authors prefer to use Fourier Transform (FT) datatypes. In [44] the magnitude and phase of 100 MHz frequency from a time-resolved signal was used for the reconstruction. A similar approach was published in [208] to retrieve the fluorescence lifetime distribution. A related idea is to directly use frequency-domain technology [55, 175, 89] to retrieve optical properties or fluorescence lifetime [207, 148]. These reconstruction approaches based on solving the inverse problem in frequency domain are a good alternative to temporal windows but are not equivalent since only MHz order frequencies are used.

In the next sections, I propose a new technique for computing fluence rate temporal windows datatypes without constraining ourselves to  $w(t) = t^n e^{-pt}$  form and preserving low computing times. I analyse different types of windows in terms of temporal selectivity, noise influence and computational efficiency. I also compare through simulations the datatypes with respect to the state-of-the-art reconstruction techniques. Finally, I reformulate the new method as a frequency based reconstruction using up to GHz order.

## 3.2. Theoretical development and analysis of new temporal windows

In this section, I will give a short review to datatypes-based reconstruction and I will describe

a new method for computing temporal windows. Further on, I will analyse different temporal windows and I will also study the noise influence on reconstruction methods.

### 3.2.1. A short review of datatype-based reconstruction

As was explained in section 2.5 of Chapter 2 time-resolved diffuse optical tomography can be posed in terms of time bins or datatypes. Absorption of the medium can be recovered by using iteratively the linearized Born approximation,

$$\delta\phi_{sd}(t) = \int_{\Omega} \left[ \phi_s^{(k)} * G_d^{(k)} \right] (\mathbf{r}, t) \delta\mu_a^{(k)}(\mathbf{r}) d\mathbf{r}, \quad (3.1)$$

where  $\delta\phi_{sd} = \phi_{sd} - \phi_{sd}^k$  is the fluence rate difference at detector  $d$  when source at position  $s$  was activated; in  $\phi_{sd}$  the experimental factors are included.  $G_d^{(k)}(\mathbf{r}, t)$  indicates the value of the Green function at every point in the domain assuming the source is located at position  $d$ .  $\phi_s^k(\mathbf{r}, t)$  is the fluence rate at every point in the domain for a source located at  $s$ .  $\delta\mu_a(\mathbf{r})$  is the absorption update at each iteration. The superscript indicates that absorption values obtained at iteration  $k$  were used; see a simplified sketch of the reconstruction procedure at Figure 3.1. Equation 3.1 can also be extended to include changes in scattering.

In the case of full-time resolved reconstruction approach (see Subsection 2.5.4) the system matrix size is  $N \cdot S \cdot D \times M$  where  $N$  is the number of time bins considered (usually between one and four thousand),  $S$  is the number of sources,  $D$  is the number of detectors and  $M$  is the number of mesh nodes. Expressing the reconstruction in terms of time has two main drawbacks: (1) the number of time bins is high (usually around one thousand), which makes the system very large and (2) it is needed to simulate the distribution of photon time of flight (DTOF) for each source and detector (which is very time-consuming).

Due to previous reasons, it is better to solve the problem of Equation 3.1 in terms of datatypes. A datatype is a filter that is applied to the DTOF to reduce the size of the problem. For example, when a temporal window filter is applied to the DTOF, the problem is reduced to

$$\begin{bmatrix} L_{sd}(w_1) \\ L_{sd}(w_2) \\ \vdots \\ L_{sd}(w_W) \end{bmatrix} = \begin{bmatrix} J_{sd}(w_1, r_1) & J_{sd}(w_1, r_2) & \cdots & J_{sd}(w_1, r_M) \\ J_{sd}(w_2, r_1) & J_{sd}(w_2, r_2) & \cdots & J_{sd}(w_2, r_M) \\ \vdots & \vdots & \cdots & \vdots \\ J_{sd}(w_W, r_1) & J_{sd}(w_W, r_2) & \cdots & J_{sd}(w_W, r_M) \end{bmatrix} \begin{bmatrix} \delta\mu_a(r_1) \\ \delta\mu_a(r_2) \\ \vdots \\ \delta\mu_a(r_M) \end{bmatrix}, \quad (3.2)$$

whose matrix size is  $W \cdot S \cdot D \times M$ , where  $W$  is the number of windows filters used. In practice, this system will be much smaller than the system of Equation 2.109 since only a few dozens of windows are usually needed.

The Fourier transform is also a datatype that converts the reconstruction problem to

$$\begin{bmatrix} L_{sd}(f_1) \\ L_{sd}(f_2) \\ \vdots \\ L_{sd}(f_F) \end{bmatrix} = \begin{bmatrix} J_{sd}(f_1, r_1) & J_{sd}(f_1, r_2) & \cdots & J_{sd}(f_1, r_M) \\ J_{sd}(f_2, r_1) & J_{sd}(f_2, r_2) & \cdots & J_{sd}(f_2, r_M) \\ \vdots & \vdots & \cdots & \vdots \\ J_{sd}(f_F, r_1) & J_{sd}(f_F, r_2) & \cdots & J_{sd}(f_F, r_M) \end{bmatrix} \begin{bmatrix} \delta\mu_a(r_1) \\ \delta\mu_a(r_2) \\ \vdots \\ \delta\mu_a(r_M) \end{bmatrix}, \quad (3.3)$$

where the matrix and the vector on the left hand side are complex. The matrix size is  $F \cdot S \cdot D \times M$ , where  $F$  is the number of frequency bins. It is interesting to notice that although

frequency based systems use frequencies in the range of MHz, in time-resolved systems, after applying the FT, the frequencies extend to the range of GHz. As it will be shown later, frequencies of GHz order provide important information for deep inclusions.

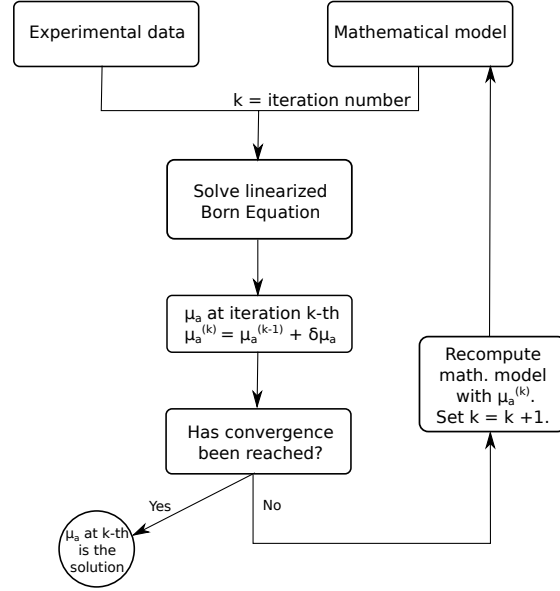


Figure 3.1: Absorption reconstruction algorithm based on linearized Born equation (Eq. 3.1).

For the following discussion, it is important to note the difference between FT- and temporal windows-based reconstruction. In the first case, the reconstruction is done directly in the frequency domain. However, in the second case, the reconstruction is performed by using temporal windows, which are proposed to be computed in the frequency domain, due to computational efficiency reasons.

### 3.2.2. A novel method to compute temporal windows

Let us define  $u(t)$  as the DTOF simulated at a detector and  $w(t)$  as an arbitrary shaped temporal window. A datatype  $\Gamma$  is defined as

$$\int_0^{\infty} u(t)w(t) dt = \Gamma, \quad (3.4)$$

since  $u(t) = 0$  for  $t < 0$ . If  $w(t) = t^n e^{-pt}$  then  $\Gamma$  can be calculated directly without computing  $u(t)$  explicitly [8]. However, until now, no method has been published that uses a different window shape without computing explicitly  $u(t)$  for each time step. The approach described in this chapter, proposes to compute the datatypes in the frequency domain; this can be done by using the Plancherel theorem as follows:

$$\Gamma = \int_{-\infty}^{\infty} u(t)\overline{w(t)} dt = \int_{-\infty}^{\infty} U(f)\overline{W(f)} df, \quad (3.5)$$

where uppercase denotes the Fourier transform defined as  $U(f) = \int_{-\infty}^{\infty} u(t)e^{-2\pi ift} dt$  and  $\overline{W(f)}$  denotes the conjugate of  $W(f)$ . If  $\overline{W(f)}$  or  $U(f)$  are non-zero at a small interval of the frequency domain, the integral of Equation 3.5 can be approximated numerically by

using small range of frequencies, see Figure 3.2. This new approach allows fast computation of a larger set of temporal windows.

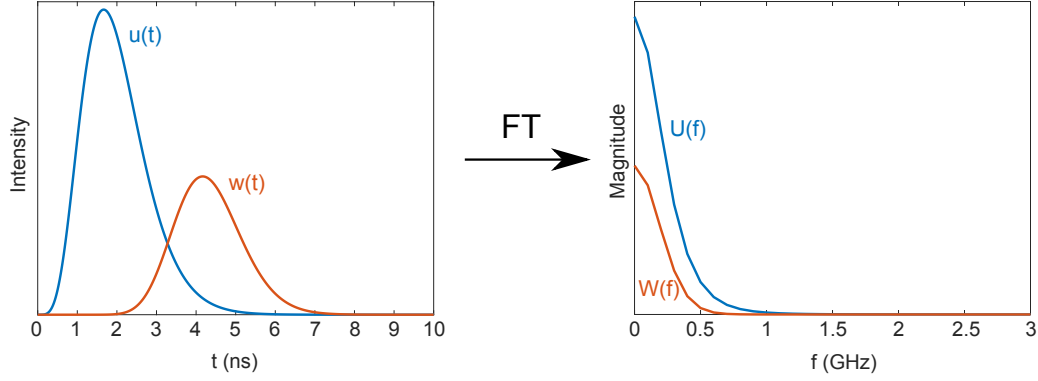


Figure 3.2: (Left) DTOF  $u(t)$  and temporal window  $w(t)$  in time domain. (Right) Magnitude of the Fourier transforms of  $u(t)$  and  $w(t)$ ; it is feasible to approximate numerically the integral in frequency domain since spectra decay rapidly.

To approximate  $\Gamma$  using few frequencies, it is important to ensure that window spectra decay rapidly. Dispersion of the windows after applying Fourier transform can be controlled by using the *uncertainty principle*. This principle states that the narrower a function  $g(t)$  is, the more spread its Fourier transform  $G(f)$  will be. That is, if one wants to have a sharp window in time domain then the integral in frequency domain will be more spread around the zero frequency. Defining the dispersion of a function as

$$D_0(g) = \frac{\int_{-\infty}^{\infty} t^2 |g(t)|^2 dt}{\int_{-\infty}^{\infty} |g(t)|^2 dt}, \quad D_0(G) = \frac{\int_{-\infty}^{\infty} f^2 |G(f)|^2 df}{\int_{-\infty}^{\infty} |G(f)|^2 df} \quad (3.6)$$

the *uncertainty principle* states that if  $g(t)$  is absolutely continuous and the functions  $tg(t)$  and  $g'(t)$  are square integrable (i.e. they are in  $L^2$  space) then [171]

$$D_0(g)D_0(G) \geq \frac{1}{16\pi^2}, \quad (3.7)$$

where  $G$  is the Fourier transform of  $g(t)$  and the equality holds only for the Gaussian density function centered at  $t = 0$ .

### Computational aspects

The reported technique has to be much faster to compute than the resolution of time-dependant Diffusion Approximation equation in order to be useful in practice. In the following section, some important computational aspects related to the time complexity will be shown.

First, to compute  $\phi(\mathbf{r}, t_n)$  in Equation 2.24, it is needed to know the previous values of  $\phi(\mathbf{r}, t_{n-i})$ , where  $t_n$  is the  $n$ -th discrete time bin and  $t_{n-i} < t_n$ . This implies that to compute  $\phi(\mathbf{r}, t_n)$  previously  $n-1$  linear systems have to be solved. However, in the frequency resolved version of Equation 2.24, there is no such restriction, since the fluence rate at any frequency can be computed independently. This is the main reason why the integral

of Equation 3.5 is faster to compute in the frequency domain. Therefore, the described technique is easily parallelizable and could be implemented in GPU hardware. Some labs have already implemented state-of-the-art photon propagation models using GPU technology [61, 192].

Another important property is that for real functions, Fourier transform coefficients will have symmetric real part and anti-symmetric imaginary part. From this follows that only positive (or negative) frequencies must be computed and only half of the integral needs to be approximated,

$$\int_{-\infty}^{\infty} U(f)\overline{W(f)} df = 2 \int_0^{\infty} U(f)\overline{W(f)} df \approx 2 \sum_{i=0}^N U(f_i)\overline{W(f_i)} \Delta f, \quad (3.8)$$

where  $N$  frequencies were used in the approximation.

However, the proposed method also suffers from one limitation. When computing temporal windows with shape  $w(t) = t^n e^{-pt}$ , the linear system to solve has the form  $\mathbf{Ax} = \mathbf{b}^{(n)} + \mathbf{x}^{(n-1)}$ , that is, as the order  $n$  of the window is increased, only the right part of the systems changes, but the matrix system is constant. This encourages to use, for example, LU factorization of the matrix  $\mathbf{A}$  to solve all the systems quickly. However, when solving the Diffusion Approximation equation at frequency domain, the systems have the form  $\mathbf{A}(f)\mathbf{x} = \mathbf{b}$  where  $f$  is the frequency. In this case, the matrix  $\mathbf{A}$  needs to be factorized for each frequency, which implies an increase in the computational cost. Nevertheless, Diffusion Approximation at frequency domain is highly parallelizable because of the independence between frequencies. Therefore, unlike for model of windows  $w(t) = t^n e^{-pt}$  the equations at each frequency-domain could be fastly solved in parallel with a GPU.

### 3.2.3. Temporal windows analysis

In the following subsections, state-of-the-art temporal windows (standard and Mellin–Laplace moments) and proposed windows for the novel method (Gaussian, Tukey and Poisson windows) are analysed in time and frequency domain. An analysis regarding their optimality for numerical integration is also included in the last subsection.

#### Standard moments

Standard moments have been used to retrieve the optical properties of homogeneous media [126] and layered models [127]. Standard moments windows are defined as  $w(t) = t^n H(t)$ , where  $n \in \mathbb{N}$  and  $H(t)$  is the Heaviside step function. These types of windows can be computed fast with state-of-the-art techniques [7]. The Fourier transform of standard moment window is

$$W(f) = \frac{n!}{(i2\pi f)^{n+1}}, \quad (3.9)$$

whose magnitude is  $n! \cdot (2\pi f)^{-(n+1)}$ . Therefore, the standard moments in time domain are equivalent to

$$\int_{-\infty}^{\infty} u(t) t^n H(t) dt = \int_0^{\infty} u(t) t^n dt = \int_{-\infty}^{\infty} U(f) \overline{(i2\pi f)^{-n-1} n!} df, \quad (3.10)$$

where the term  $(i2\pi f)^{-n-1}n!$  can be considered as a window in frequency domain. This window magnitude tends to a Dirac delta distribution as  $n$  goes to infinity, which is obvious since  $t^n H(t)$  will approach a step function with infinity value at  $t \geq 0$  and the only non-zero magnitude frequency in the spectrum will be the zero frequency. Therefore, the higher the order, the smaller the frequency range covered by the windows.

If the standard moments are centralized by the time of flight  $\langle t \rangle = \int_{-\infty}^{\infty} u(t) t H(t) dt$  then

$$\begin{aligned} \int_{-\infty}^{\infty} u(t) (t - \langle t \rangle)^n H(t) dt &= \int_0^{\infty} u(t) (t - \langle t \rangle)^n dt \\ &= \int_0^{\infty} u(t) \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} t^k \langle t \rangle^{n-k} dt \\ &= \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} \langle t \rangle^{n-k} \int_0^{\infty} u(t) t^k dt, \end{aligned} \quad (3.11)$$

which is a weighted sum of standard moments. Now, I introduce the state-of-the-art windows known as Mellin–Laplace moments.

### Mellin–Laplace moments

Mellin–Laplace windows are defined as  $w(t) = t^n e^{-pt} H(t)$  where  $p > 0$  and  $n \in \mathbb{N}$ . These windows can also be computed fast [101]. Their Fourier transform is given by

$$W(f) = \frac{n!}{(p + i2\pi f)^{n+1}}. \quad (3.12)$$

In Figure 3.3 and 3.4, the windows for different  $n$  orders and  $p$  values are shown. The higher the Mellin–Laplace order  $n$  the narrower the window magnitude will be in frequency domain. As expected from the uncertainty principle, the higher the  $p$  value the more spread the window will be in frequency domain.

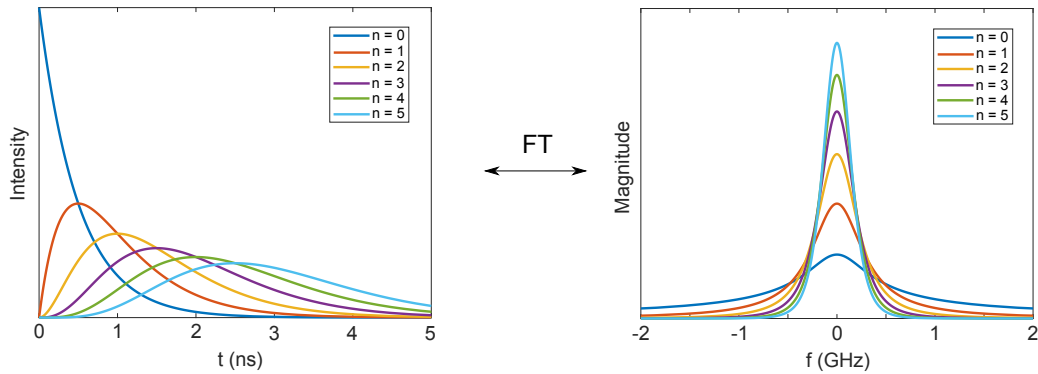


Figure 3.3: (Left) Different  $n$  orders of Mellin–Laplace windows for  $p = 2$ . (Right) Magnitude of the Fourier transform of those windows; as the order  $n$  of the Mellin–Laplace moment is increased, its spectrum magnitude is narrower and has less dispersion.



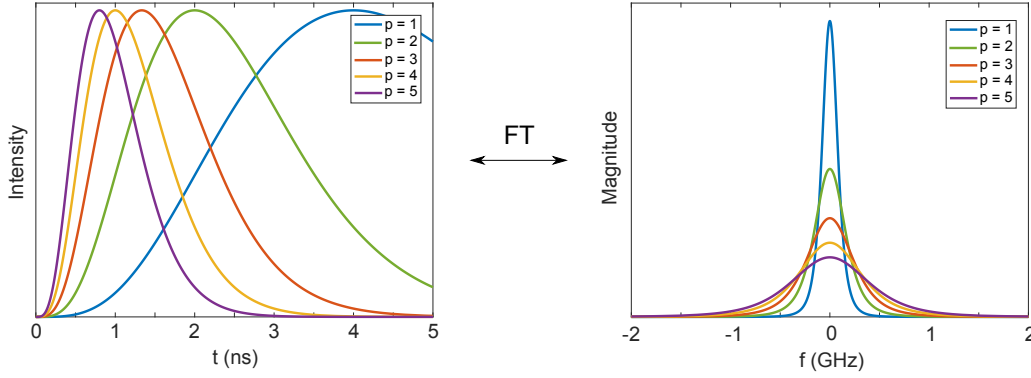


Figure 3.4: (Left) Fifth order ( $n = 4$ ) Mellin–Laplace window with different  $p$  values. (Right) Magnitude of the Fourier transform of those windows; as the  $p$  value is increased, the magnitude of the Mellin–Laplace window spectrum is more spread.

Zero order Mellin–Laplace datatypes are Laplace transforms for different  $p$  values. Recently, an optical system was proposed [95], which computes Laplace datatypes without the need of measuring the complete DTOF. The system described by Hasnain et al. [95] is quite interesting, since it is much faster than typical time-resolved systems. Nevertheless, Laplace datatypes have poor temporal selectivity (they cover a broad range of time) and are highly correlated, which does not make them the best candidates in terms of tomography capabilities.

### Generalized Gaussian window

The Gaussian function and its Fourier transform are

$$w(t) = e^{-t^2/(2\sigma^2)} \longleftrightarrow W(f) = \sigma\sqrt{2\pi}e^{-2\sigma^2\pi^2 f^2}. \quad (3.13)$$

As expected from the uncertainty principle, the higher  $\sigma$ , the more spread the window at time domain will be and, therefore, the narrower it will be in the frequency domain, see Figure 3.5. Note that the centralized Gaussian windows gives the lower bound of the uncertainty principle.

The Laplacian or exponential window (see Figure 3.6) are defined as

$$w(t) = \exp(-p|t|) \longleftrightarrow W(f) = \frac{2p}{p^2 + (2\pi f)^2}, \quad (3.14)$$

where  $p$  is a parameter.

Both Gaussian and Laplacian windows can be described by the generalized Gaussian window (also known as the generalized normal window),

$$w(t) = \frac{\beta}{2\sigma\Gamma(1/\beta)} e^{-(|t|/\sigma)^\beta} \longleftrightarrow W(f) = \Gamma^{-1}\left(\frac{1}{\beta}\right) \sum_{n=0}^{\infty} \frac{(-1)^n \pi^{2n} f^{2n}}{(2n)!} \sigma^{2n} \Gamma\left(\frac{2n+1}{\beta}\right). \quad (3.15)$$

where  $\beta$  is a free parameter and  $\Gamma$  is the gamma function. If  $\beta = 2$  the function is Gaussian and if  $\beta = 1$  is Laplacian. As  $\beta \rightarrow \infty$  the window converges to a square window.

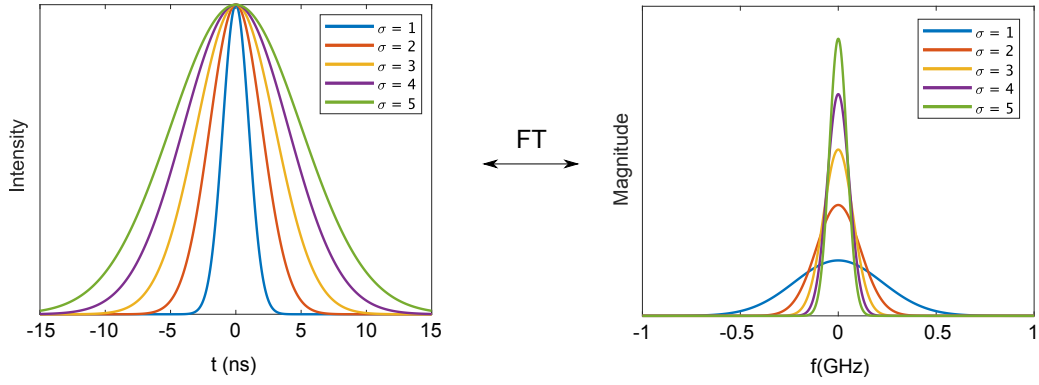


Figure 3.5: (Left) Several Gaussian functions with different  $\sigma$  values. (Right) Spectrum magnitude of previous Gaussian functions. The more spread the Gaussian function, the narrower its Fourier transform magnitude.

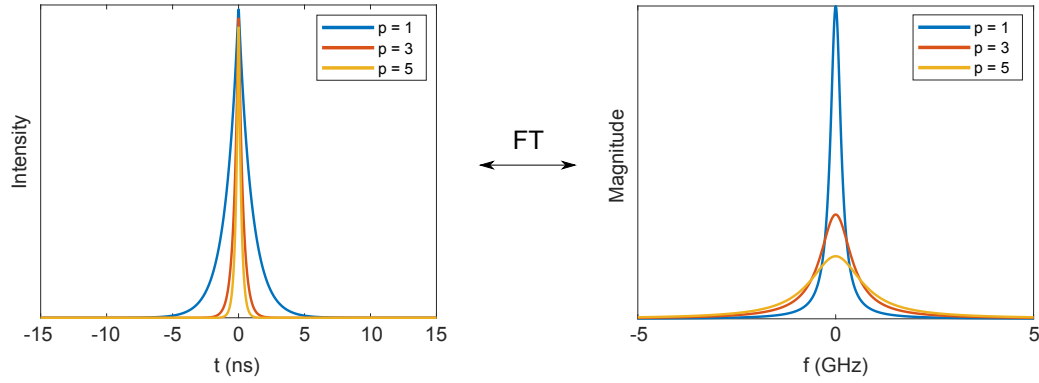


Figure 3.6: (Left) Several Laplacian windows with different  $p$  values. (Right) Spectrum magnitude of previous functions.

### Tukey window

The centralized Tukey window [94] is defined as

$$w(t) = \begin{cases} 1, & 0 \leq |t| \leq \alpha t^* \\ 0.5 \left[ 1 + \cos \left( \pi \frac{t - \text{sign}(t) \alpha t^*}{\text{sign}(t) (t^* - \alpha t^*)} \right) \right], & \alpha t^* \leq |t| \leq t^* \\ 0, & \text{elsewhere,} \end{cases} \quad (3.16)$$

where  $t^*$  is the limit of the window and  $0 \leq \alpha \leq 1$  parameter controls the smoothness of the window.

If  $\alpha = 1$  the Tukey window is rectangular and its Fourier transform is defined as

$$w(t) = \text{rect}(t/\tau) \longleftrightarrow W(f) = \tau \text{sinc}(\pi f \tau). \quad (3.17)$$

These windows have a high selectivity of photon arrival time. The spectrum magnitude of rectangular function ( $\alpha = 1$ ) is a sinc function with infinite oscillations, which can make

difficult to integrate them numerically. Nevertheless, as the  $\alpha$  value decreases, the oscillations at high frequencies disappear without losing much temporal selectivity.

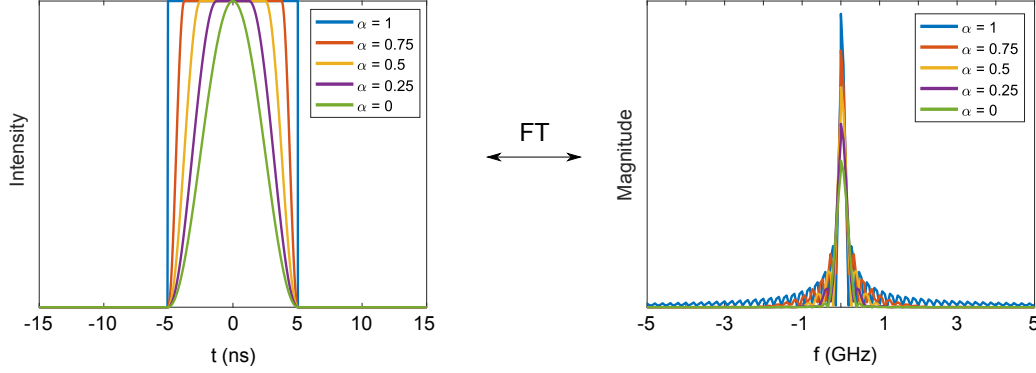


Figure 3.7: (Left) Several Tukey functions with different  $\alpha$  values. (Right) Magnitude of the Fourier transforms of Tukey functions.

### Trade-off between temporal selectivity and computation complexity

An optimal temporal window is the one that is easy to compute (it has few oscillations and small dispersion at frequency domain) and at the same time its temporal selectivity is good enough. For some windows, optimality can be described by  $D_0^* = D_0(g)D_0(G)$  (see Subsection 3.2.2). The  $D_0^*$  of a window quantifies the trade-off between the temporal selectivity and easiness to compute as expressed by uncertainty principle. The lower the value  $D_0^*$  is for a given window, the better the window temporal selectivity will be and the less frequencies will be needed to perform the numerical integration.

In Appendix F the dispersion of each window was calculated. In Table 3.1, the main optimality results were summarized.

| Window         | $D_0^*$  |
|----------------|--|
| Mellin–Laplace | $\frac{1}{16\pi^2} \frac{(2n+2)!}{(2n)!} \left( \frac{2n}{2n-1} - 1 \right), \quad (n \geq 1)$ |
| Gaussian       | $1/(16\pi^2)$  |
| Exponential    | $1/(8\pi^2)$   |

Table 3.1:  $D_0^*$  for each analyzed window. Standard moments and Tukey windows are not  $L^2$  integrable.

The Gaussian is the window that has the lowest  $D_0^*$  value; this was expected from the uncertainty principle theorem and this is what makes the Gaussian window one of the most promising windows. The exponential window has a twice bigger value compared to the Gaussian window; this means that the numerical integral could be a bit more difficult to compute but it is still a good candidate. Mellin–Laplace windows will increase its  $D_0^*$  value almost linearly for increasing orders although it is independent of  $p$  value. Expression  $D_0^*$  is not in  $L^2$  space for Tukey windows, nevertheless from Figure 3.7 it is evident that for some  $\alpha$  values is still feasible to perform numerical integration at frequency domain. Standard

moments are neither  $L^2$  integrable, but since by nature they are highly correlated they will not be used in this work.

A point that should also be considered is the shift theorem of the Fourier transform. It states that given a function  $g(t)$ , if this function is shifted an amount of  $t_0$ , then its Fourier transform is

$$G_{t-t_0}(f) = G(f)e^{-i2\pi ft_0}, \quad (3.18)$$

where  $G_{t-t_0}(f)$  indicates the Fourier transform of shifted function  $g(t-t_0) = g(t) * \delta(t-t_0)$  and the complex exponential does not affect the spectral magnitude, since  $|e^{-i2\pi ft_0}| = 1$ , but it affects the phase. This theorem implies that as temporal windows are shifted to later times, more oscillations will appear. However, in practice this phenomenon did not prevent us to perform the numerical integration accurately.

The new method to compute temporal windows is also useful for windows with the form  $w(t) = t^n e^{-pt}$ , such as Mellin–Laplace or standard moments. Both of these moments suffer from an important handicap if traditional computation methods are used: in order to compute the  $n$ -th order window then all the previous orders,  $1, \dots, n-1$ , must be computed (in the case of Mellin–Laplace if  $p$  parameter is modified then all orders must be computed again). These problem can be critical in the case of Mellin–Laplace since as the order of the moments increase, the window temporal selectivity decreases (see Figure 3.3 (left)). If it is wanted to have better sensitivity to later photons (i.e. photons that probabilistically have go deeper into the tissue) then higher  $p$  values must be used. There are two ways of doing it: the first option is to use the same  $p$  value for all windows; but the main drawback is that many orders will have to be calculated, since a higher  $p$  value shifts windows average time-of-flight to earlier photons (see Figure 3.4). The second option implies to use windows with different  $p$  values (e.g. a low  $p$  value for early windows and high  $p$  values for late windows) which is not computationally efficient since if the  $p$  value is changed then all orders up to  $n$  will have to be computed again as explained before. Nevertheless, the new approach presented in this chapter does not suffer from these bottlenecks since Mellin–Laplace moments can be computed for each  $p$  and  $n$  value independently (same for  $n$  order in the case of standard moments). Therefore, one of the biggest advantages of the presented method is not only that new windows can be computed but also that Mellin–Laplace windows for different  $p$  values can be computed more efficiently.

### 3.2.4. Noise influence

Since DTOF signals are transformed to datatypes space, the noise will also suffer changes. In this part, noise transformations and correlations are analysed to understand noise influence.

The DTOF signal measured by a detector is theoretically described as

$$u(t) = x(t) + \varepsilon(x(t)), \quad (3.19)$$

where  $x(t)$  is the noiseless signal and  $\varepsilon$  is the noisy term which depends on  $x(t)$  magnitude, uncorrelated (that is,  $\text{Cov}(\varepsilon(x(t)), \varepsilon(x(t'))) = 0$  for  $t \neq t'$ ) and follows a Poisson distribution. Applying the Fourier transform to  $u(t)$  yields,

$$\begin{aligned} U(f) &= X(f) + \Upsilon(f) \\ &= \int_{-\infty}^{\infty} x(t)e^{-i2\pi ft} dt + \int_{-\infty}^{\infty} \varepsilon(x(t))e^{-i2\pi ft} dt, \end{aligned} \quad (3.20)$$

where the second term in the right part is the noisy contribution and  $U(f)$  can be considered a random variable. As stated before, in time domain, the noise at each time bin is independent, however after applying the Fourier transform this property does not hold anymore. For example, the covariance between two different frequencies is,

$$\begin{aligned}
E \left[ U(f_1) \cdot \overline{U(f_2)} \right] - E[U(f_1)] \cdot E \left[ \overline{U(f_2)} \right] &= E \left[ \int_{-\infty}^{\infty} u(t) e^{-i2\pi f_1 t} dt \cdot \overline{\int_{-\infty}^{\infty} u(t) e^{-i2\pi f_2 t} dt} \right] \\
&\quad - E \left[ \int_{-\infty}^{\infty} u(t) e^{-i2\pi f_1 t} dt \right] \cdot E \left[ \overline{\int_{-\infty}^{\infty} u(t) e^{-i2\pi f_2 t} dt} \right] \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (E[u(t) \cdot u(t')] - E[u(t)] \cdot E[u(t')]) \\
&\quad e^{-i2\pi f_1 t} e^{i2\pi f_2 t'} dt dt' \\
\text{(Due to TR signal noise independence)} &= \int_{-\infty}^{\infty} E[u(t)] e^{-i2\pi(f_1 - f_2)t} dt \\
(E[u(t)] = x(t), \text{ due to shot/Poisson noise}) &= \int_{-\infty}^{\infty} x(t) e^{-i2\pi(f_1 - f_2)t} dt \\
&= X(f_1 - f_2),
\end{aligned} \tag{3.21}$$

so the covariance of Fourier transform signal at two different frequencies is equal to the Fourier coefficient of the noiseless signal at frequency  $f_1 - f_2$ . An alternative derivation can be seen at Appendix G.

From this result, the covariance between any temporal window datatypes can be determined. Let us define a window datatype shifted  $kt_0$  in time as,

$$M_{k,t_0} = \int_0^{\infty} u(t) w(t - kt_0) dt = \int_{-\infty}^{\infty} U(f) \overline{W(f)} e^{i2\pi f kt_0} df, \tag{3.22}$$

where  $t_0$  is the shift unit length and  $k$  is the number of shifts steps that are taken. Then, the covariance between shifted window datatypes is,

$$\begin{aligned}
\text{Cov}(M_k, M_l) &= E \left[ \int_{-\infty}^{\infty} U(f) \overline{W(f)} e^{i2\pi f kt_0} df \cdot \overline{\int_{-\infty}^{\infty} U(f) \overline{W(f)} e^{i2\pi f lt_0} df} \right] \\
&\quad - E \left[ \int_{-\infty}^{\infty} U(f) \overline{W(f)} e^{i2\pi f kt_0} df \right] \cdot E \left[ \overline{\int_{-\infty}^{\infty} U(f) \overline{W(f)} e^{i2\pi f lt_0} df} \right] \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left( E[U(f) \cdot \overline{U(f')}] - E[U(f)] \cdot E[\overline{U(f')}] \right) \\
&\quad \overline{W(f)} e^{i2\pi f kt_0} W(f') e^{-i2\pi f' lt_0} df df' \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} X(f - f') \overline{W(f)} W(f') e^{i2\pi(k t_0 f - l t_0 f')} df df'.
\end{aligned} \tag{3.23}$$

From Equation 3.23, the covariance and correlation matrices of any window can be obtained. Note that for standard (and Mellin–Laplace) moments the shift is already fixed by  $n$  (and  $p$ ) value. In Figure 3.8, an example of the typical correlation matrices of Mellin–Laplace,

Gaussian and Tukey windows are shown. It is evident that, for a fixed  $p$  value, as the order  $n$  of Mellin–Laplace window increases, its get more and more correlated with neighbour windows (see Figure 3.3 to get an intuitive idea). However, for Gaussian and Tukey windows, the correlation can be minimized by separating them far enough; in the example given the Gaussian windows are not correlated with any other windows, because the windows did not overlap. The same thing could have been done with Tukey windows.

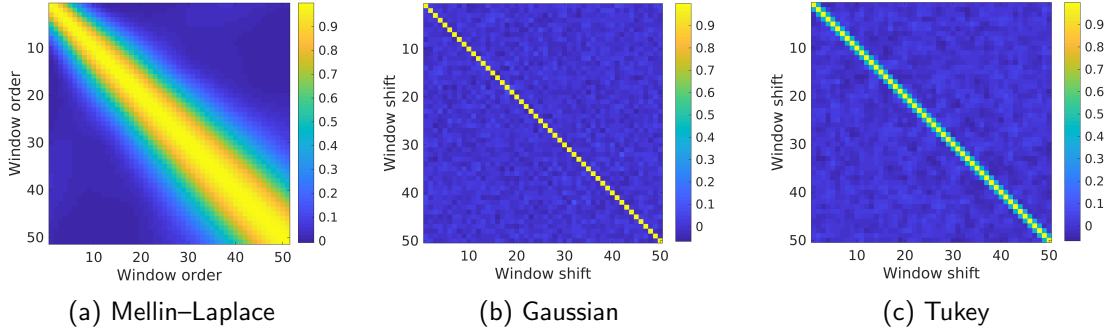


Figure 3.8: Correlation matrices of Mellin–Laplace (first 50 orders,  $p = 7$ ), Gaussian ( $t_0 = 0.16$  ns shifts,  $\sigma = 0.05$ ), and Tukey windows ( $t_0 = 0.16$  ns shifts,  $\alpha = 0.25$  and  $t^* = 0.2$  ns).

In time-domain reconstruction, the heteroscedasticity [14] is present in the noise, that is, the noise at each time bin has different variance but they are not correlated with each other (this is due to photon noise physics). After applying overlapping windows, such as Mellin–Laplace or Fourier transform, non-zero covariance values will appear. Therefore, in these cases, Least Squares (LS) will not be a Best Linear Unbiased Estimator (BLUE) anymore. Nevertheless, Generalized Least Squares (GLS) (also known as Aitken’s estimator) can be used instead, which is also a BLUE [14]. In theory, LS and GLS will reach the same solution for different noise realizations, however since for GLS the inverse of covariance matrix must be computed, some problems could arise if it is ill-posed. So, noise correlations should be avoided whenever possible.

### 3.3. Methods

The setup of the simulations, optical properties and numerical parameters are presented.

#### 3.3.1. Numerical simulations

To analyze the reconstruction improvements that could be obtained with the decorrelated windows, some numerical simulations were performed. The used computational phantom was a three-dimensional volume with a spherical inclusion at different depths. This phantom can be consider as an approximation to functional near-infrared spectroscopy experiments where the inclusion represents the activation in the cortex. The computational phantom had a size of  $9 \times 9 \times 5$  cm<sup>3</sup> and was discretized using around 33 and 180 thousand nodes and elements respectively. Numerical simulations were performed in a reflectance geometry by solving the time-domain Diffusion Approximation up to 10 ns; the boundary conditions were

implemented as described in [195]. Each simulated curve contained up to  $2 \times 10^5$  photons, were convoluted with the instrumental response function of a single photon avalanche diode detector (SPAD, FWHM  $\approx 160$  ps) and corrupted by Poisson noise, see Figure 3.9.

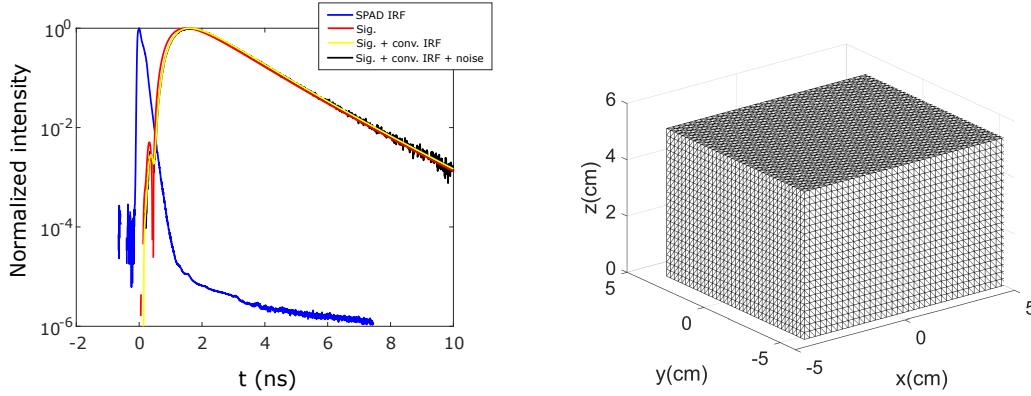


Figure 3.9: (Left) Instrumental response function of a SPAD detector (blue line). Normalized simulated signal without IRF neither noise (red line). Normalized simulated signal convoluted with the IRF (yellow line). Normalized simulated signal convoluted with IRF and corrupted with Poisson noise (black line). (Right) Tetrahedral mesh used for simulations; it had a size of  $9 \times 9 \times 5$  cm<sup>3</sup> and was discretized using around 33 and 180 thousand nodes and elements respectively.

In this work, only absorption inhomogeneities were considered, although it can be easily extended to scattering inhomogeneities. Background absorption was set to  $\mu_a = 0.018$  cm<sup>-1</sup> and the reduced scattering was  $\mu'_s = 14.7$  cm<sup>-1</sup> over all the domain. An inclusion of 0.5 cm radius was included whose optical properties were  $\mu_a = 0.337$  cm<sup>-1</sup> and  $\mu'_s = 14.7$  cm<sup>-1</sup> (these values were taken from the experiment performed at [257] using 550 nm wavelength light and are of similar order as found at several types of human tissue).

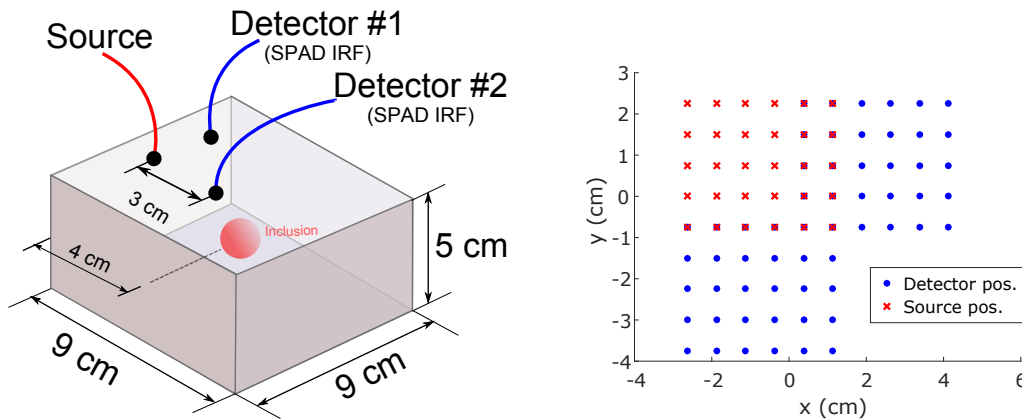


Figure 3.10: (Left) Scheme of the numerical phantom. Simulated signals were convoluted with the IRF of a SPAD and corrupted by Poisson noise. (Right) Source positions; red crosses. Detector positions; blue dots. Note that for each source only the signal at two detectors were considered; 3 cm right and 3 cm down. For example, for source at position  $(-2.62, 2.25)$  cm only the signal at detectors  $\mathbf{x}_1 = (0.38, 2.25)$  cm and  $\mathbf{x}_2 = (-2.62, -0.75)$  cm was used.

The setup of source and detectors was moved along the top boundary to scan the phantom at multiple positions. The distance between the source and detectors was 3 cm at orthogonal directions, see Figure 3.10. The scan of the computational phantom with the inclusion was done at 30 different source positions (6 shifts in  $x$ -direction separated by 0.75 cm and 5 shifts in  $y$ -direction separated by 0.75 cm). The inclusions were set from 2 to 3.5 cm depth.

During the reconstruction, to avoid large sensitivities in the surface nodes the sensitivity of each node was normalized by a diagonal weight matrix with values proportional to nodes depth. A flowchart of the used algorithm is shown at Figure 3.11. Simulations and reconstructions were done using a laptop with an Intel Core i7 processor and 4 GB RAM.

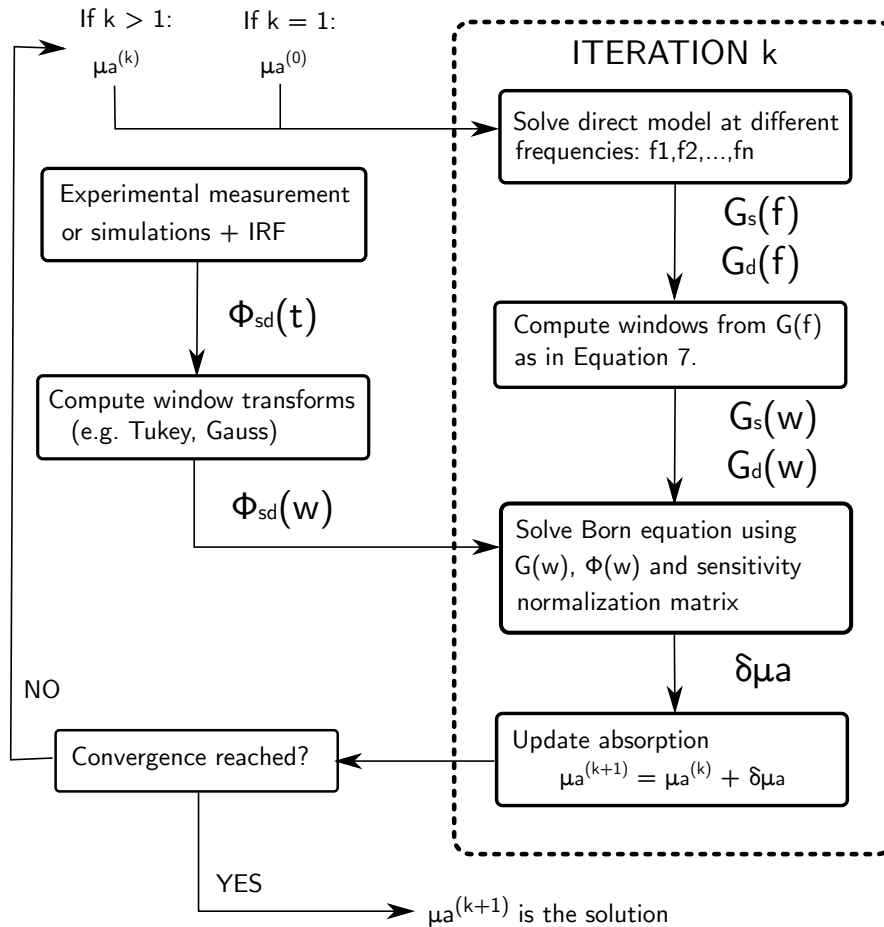


Figure 3.11: Reconstruction algorithm flowchart.

### 3.3.2. Quantitative evaluation metrics

Three different evaluation metrics (localization error, average contrast and relative recovered volume) were used to measure the reconstructions quality from different points of view.

Localization error is defined as Euclidean distance between the center of mass of simulated inclusion  $x_s$  and the center of mass of the reconstructed inclusion  $x_r$ . If the reconstructed optical properties are identical to the truth, the localization error is zero. The recovered inclusion is determined by selecting the reconstructed absorption changes that are greater



than a given threshold defined later.

$$\text{Localization error} = \|\mathbf{x}_s - \mathbf{x}_r\|_2. \quad (3.24)$$

The average contrast evaluation metric is based on the mean value of the region of interest (i.e. inclusion volume):

$$\text{Average contrast} = \frac{\sum_{i \in N_r} \mu_i}{|N_r| \hat{\mu}} \quad (3.25)$$

where  $\mu_i$  denotes the reconstructed absorption at node  $i$ ,  $N_r$  represents the set of nodes at the region of interest,  $|N_r|$  denotes the number of nodes within the set and  $\hat{\mu}$  is the truth values of absorption at the region of interest. If the reconstructed absorption is identical to the truth, the average contrast value is one.

The last evaluation metric is the relative reconstructed volume ( $V_{RRV}$ ) which is defined as

$$V_{RRV} = V_r/V_s \times 100\% \quad (3.26)$$

where  $V_r$  and  $V_s$  are the volume of the reconstructed inclusion and simulated case respectively. The volume of the reconstructed inclusion,  $V_r$ , is computed by thresholding the recovered absorption changes and computing the volume of those elements. If the reconstructed absorption is identical to the truth, the relative reconstructed volume is one.

## 3.4. Results and discussion

In this section, the simulations results are given in order to analyze the performance of tomography algorithms based on temporal windows and Fourier transform datatypes. Subsequently, different perspectives of the same problem are given and results are discussed.

### 3.4.1. Comparing state-of-the-art windows with Tukey and Gaussian windows

The purpose of the first simulations is to check whether Tukey or Gaussian windows improve the reconstruction in depth compared to  $w(t) = t^n e^{-pt}$  state-of-the-art windows.

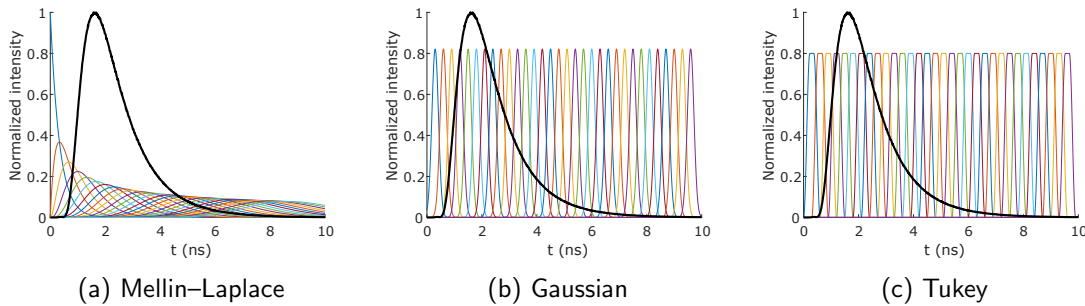


Figure 3.12: Black curve represents a simulated DTOF. (Left) Mellin-Laplace windows ( $p = 3$ , 35 moments), note that they are highly overlapped. (Middle) Gaussian windows ( $\sigma = 0.3$ ). (Right) Tukey windows ( $\alpha = 0.25$ ,  $t^* = 0.3$  ns).

For the Mellin–Laplace reconstruction  $p = 3$  and the first 35 moments were used (Figure 3.12 Left). For Gaussian windows, the value of  $\sigma$  was set to 0.3 and their centers were located from 0.3 to 9.6 ns in steps of 0.3 ns (32 windows in total, Figure 3.12 Middle). For the Tukey windows, the same centers were used and the parameters were set to  $\alpha = 0.25$  and  $t^* = 0.3$  ns (Figure 3.12 Right). The given parameters were selected so that in all cases the windows totally covered the simulated signal and the number of windows were similar (35 for Mellin–Laplace and 32 for Tukey and Gaussian). The computed frequencies ranged from 0 to 2 GHz in steps of 200 MHz.

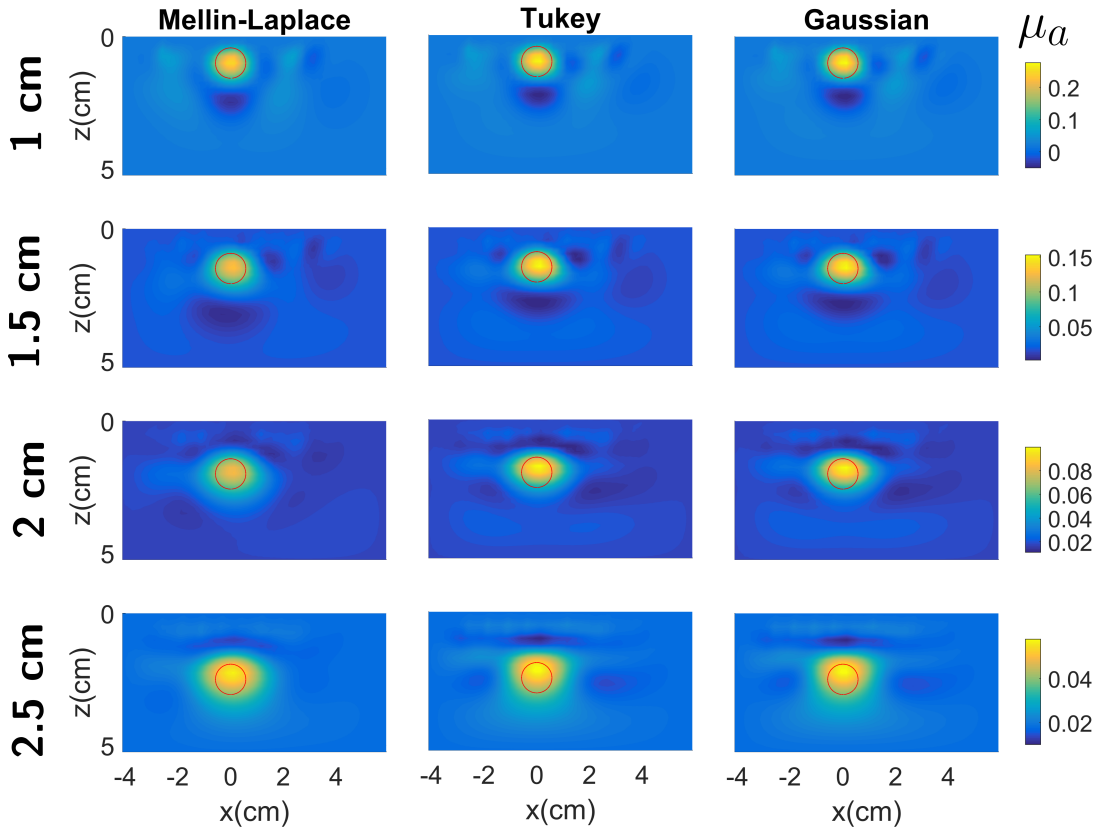


Figure 3.13: Absorption,  $\mu_a$ , reconstructions using Mellin–Laplace, Tukey and Gaussian windows for inclusions from 1 cm to 2.5 cm deep. The red circle indicates the correct location of the inclusion.

The Figure 3.13 shows reconstruction using Melin-Laplace, Tukey and Gaussian temporal windows for inclusions that ranged from 1 to 2.5 cm deep (in steps of 0.5 cm). In Figure 3.14, the reconstructions for inclusions from 3 cm to 4 cm deep is given. In Figure 3.15, the evaluation metrics results are shown for each method. For inclusions that are 3 cm or shallower the reconstruction in terms of localization does not depend significantly on the types of windows that are used since localization error is below 0.25 cm for all windows. However, absorption quantification is constantly underestimated by Mellin-Laplace windows, see how average contrast metric decreases exponentially and the contrast of Tukey and Gaussian windows is always greater than state-of-the-art Mellin-Laplace windows. For an inclusion that is 3.5 cm depth Mellin–Laplace windows reconstruction is significantly worse in terms of inclusion localization; note that localization error for Mellin-Laplace windows

dramatically increases getting bigger than 1.5 cm for inclusions 3.5 cm deep; these results are in agreement with [181] which shows the same problem beyond 2 cm depth. Moreover, the relative reconstructed volume also increases due to the fact that Mellin-Laplace loses the inclusion localization and absorption quantification, and therefore, it tries to compensate it by reconstructing higher volume inclusions. This phenomenon occurs due to the diffusion of photons inside the medium. In contrast, the results also show that Tukey and Gaussian windows give very similar reconstructions and their localization error is always less than 0.5 cm.

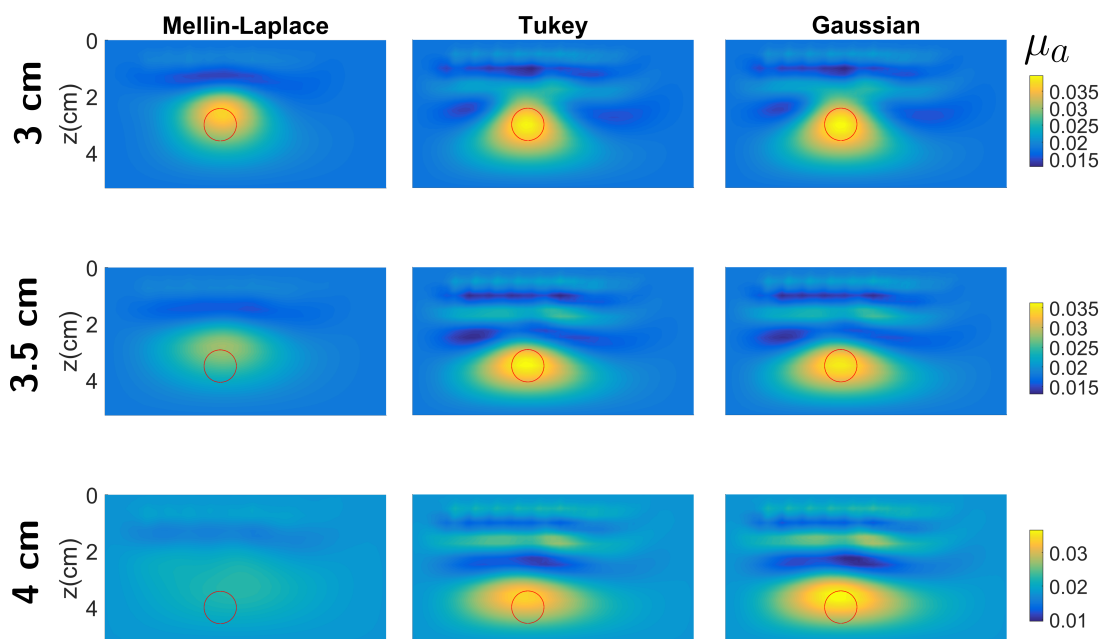


Figure 3.14: Absorption,  $\mu_a$ , reconstructions using Mellin-Laplace, Tukey and Gaussian windows for inclusions from 3 cm to 4 cm deep. The red circle indicates the correct location of the inclusion.

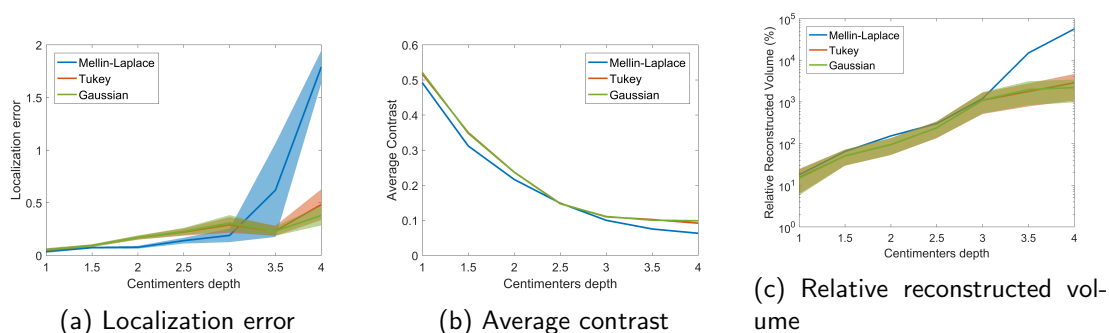
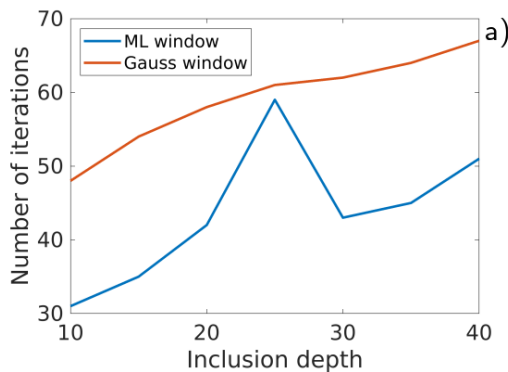


Figure 3.15: Mellin-Laplace (blue), Tukey (red) and Gaussian (yellow). Quantitative evaluation metric values for each window and inclusion depth. For localization error and relative reconstructed volume, the standard error shadows indicate the uncertainty given by selecting interest regions whose thresholds range from 60% to 80% of maximum absorption.

The fact that proposed windows outperform Mellin-Laplace state-of-the-art windows for deep

inclusions can be explained by the theoretical analysis shown in the previous sections. First, Mellin–Laplace windows have a poor late-arrival photon selectivity and windows are highly correlated. This implies that each Mellin–Laplace window has very little new information by itself. As the order of the Mellin–Laplace window increases the new information that it carries decreases. That implies that a huge amount of window should be used to include all the available information in the DTOFs. This evidence demonstrate the superior performance of Gaussian and Tukey window compared to Mellin–Laplace because of the better photon time arrival selectivity and less noise correlation.

Regarding the computation time of the reconstructions, in Table 2 the number of seconds per iteration is given for each method. To compute windows with form  $t^n e^{-pt}$  is much faster than the proposed approach and full-time reconstruction. Specifically,  $t^n e^{-pt}$  windows are around eight times faster to compute than the proposed approach because  $LU$  factorization can be used. However, for the proposed method factorization is not efficient since the matrix changes at each frequency. Full-time resolved approach is nineteen times slower than  $t^n e^{-pt}$  windows and more than twice slower than the proposed approach. These results shows the trade-off of the proposed method: more information in depth is obtained but computation time is larger than  $t^n e^{-pt}$  windows. However, it should also be taken into account that novel method approach is highly parallelizable, that is, the Diffusion Approximation equation can be solve independently for each frequency. Note that this property does not hold for the direct computation of  $t^n e^{-pt}$  windows [7] and full-time approach. Therefore, the computation time of novel approach should considerably improve with GPU usage due to its intrinsic parallelizable nature. In Figure 3.16(a), the number of computed iterations until convergence is given. Convergence was reached when relative error  $\delta\mu_a/\mu_a^{(k)} < 5 \times 10^{-3}$ . It can be seen that Mellin–Laplace windows converge faster mainly because it underestimates absorption quantification.



b)

| Window:    | $t^n e^{-pt}$ | Novel approach | Full-time      |
|------------|---------------|----------------|----------------|
| secs/iter. | $47 \pm 3$ s  | $380 \pm 10$ s | $875 \pm 10$ s |

Figure 3.16 & Table 2: (a) Number of iterations until convergence for Mellin–Laplace and Gaussian windows. (b) Seconds per iteration for Mellin–Laplace windows, novel approach (with Gaussian) and full-time resolved approach.

As explained before, the simulated DTOFs were convolved with an SPAD IRF and corrupted with photon noise. These DTOFs were deconvolved with a Wiener filter before performing the reconstruction. In some papers, such as [101], a different approach is proposed: DTOFs from a reference medium  $A$  and objective medium  $B$  are used. The optical properties of medium  $A$  are already known because an homogeneous medium is taken as reference and it includes implicitly all the information regarding instrumental factors. The medium  $B$  is the

medium to be reconstructed. Absorption of medium  $B$  can be recovered by using iteratively the linearized cross-Born equation,

$$G_{sd}^A(\mathbf{r}, t) * M_{sd}^B(\mathbf{r}, t) - M_{sd}^A(\mathbf{r}, t) * G_{sd}^{B(k)}(\mathbf{r}, t) = -M_{sd}^A(\mathbf{r}, t) * \int_{\Omega} \left[ G_s^{B(k)} * G_d^{B(k)} \right] (\mathbf{r}, t) \delta\mu_a \, d\mathbf{r}, \quad (3.27)$$

where  $*$  is the convolution operator,  $M_{sd}^A$  and  $M_{sd}^B$  indicate the measurement obtained at reference and to be reconstructed media by detector  $d$  when source  $s$  was activated,  $k$  indicates the iteration number,  $G_{sd}^A$  and  $G_{sd}^{B(k)}$  is the simulated Green's function (free of noise and instrumental factors) from source  $s$  at detector  $d$  by using the photon propagation mathematical model described in Equation 2.24 and  $G_s^{B(k)}$  indicates the Green's function value at every point in the domain given by the propagation model.

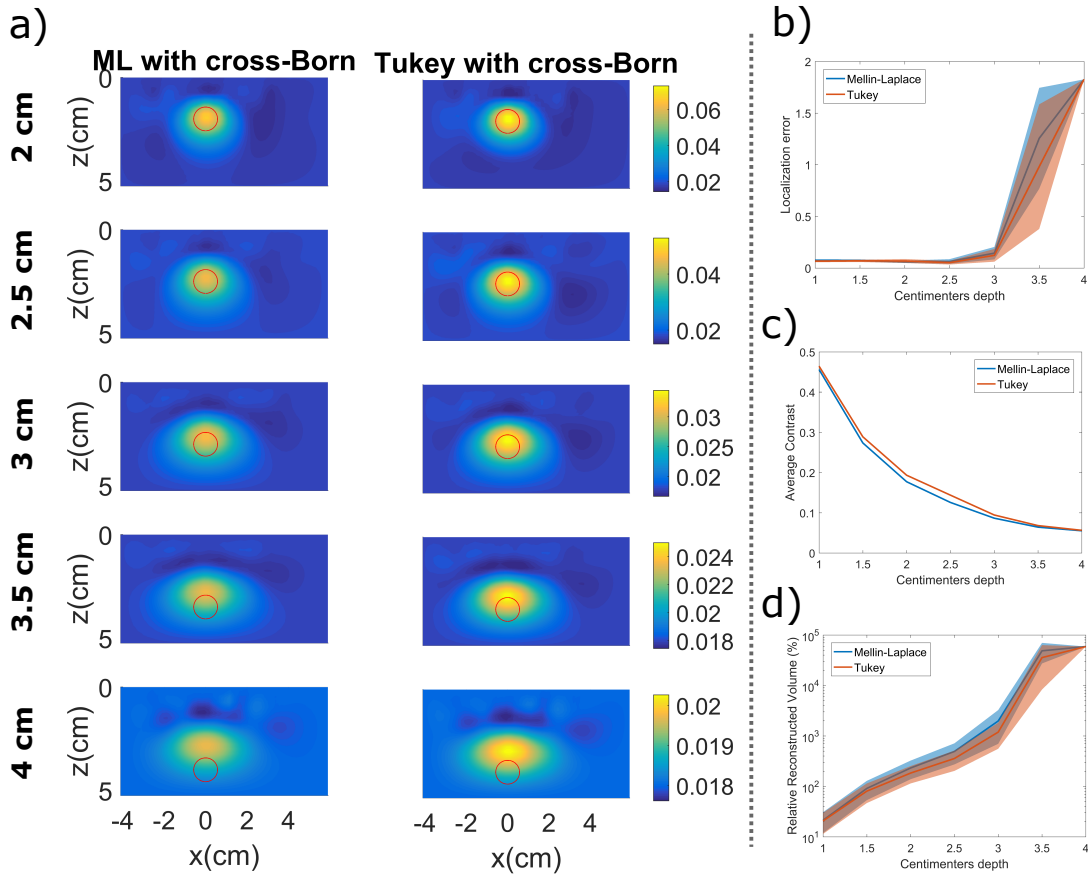


Figure 3.17: (a) Reconstructions for Mellin-Laplace and Tukey windows from same simulated data as 3.13; cross-Born approach was used to deal with instrumental factors such as IRF or photon noise. The red circle indicates the correct location of the inclusion. (b,c,d) Quantitative metrics for cross-Born based reconstructions.

The cross-Born equation offers the ease of not having to deal with instrumental factors directly (no deconvolution of the measurements is needed). However, it also implies to convolve the data and lose some information in the process. In Figure 3.17 (a), reconstructions using Mellin-Laplace and Tukey windows with the cross-Born approach are shown. We chose Tukey windows since they give very similar results to Gaussian windows. In Figure 3.17 (b,c,d), the quantitative metrics for those simulations are given. Results for inclusions as

deep as 3 cm are very similar to the deconvolution approach, but for inclusion deeper or equal to 3.5 cm the localisation accuracy is lost, see Figure 3.17 (b). Moreover, the contrast is lower than previous reconstructions; although, slight differences are seen between Mellin-Laplace and Tukey windows even using cross-Born equation, see Figure 3.17 (c). The relative recovered volume increases considerably for both methods (Figure 3.17, d). These results show that when the cross-Born approach is used, some information in depth is lost for inclusions deeper than 3 cm and to take full advantage of not overlapping datatypes (e.g. Tukey or Gaussian windows), experimental data must be deconvoluted before performing a reconstruction.

From a different point of view, this theoretical analysis can also be viewed as an explanation of full time-resolved tomography limits. That is, Tukey windows-based reconstruction can be seen as an intermediate step between correlated windows-based reconstruction and full time-resolved reconstruction (i.e. reconstruction based on using each time bin as a datatype). In fact, in the last years several labs [61, 192], have developed photon propagation models for Graphics Processing Units (GPU). One of the goals of using GPU is to compute the full DTOF curve and to fit it entirely as described in Equation 2.109. To fit the entire DTOF curve is the limiting case of very narrow Tukey windows; if windows are shrunk to time bin size it will converge to DTOF curve fitting. Therefore, some of the questions that arise from this theoretical work is, how much benefit can be obtained by using full time-resolved reconstructions? Does the computational cost of computing the full curve outweighs the gain of some information? Moreover, since the computation of decorrelated windows, such as Tukey or Gaussian, are easily parallelizable and adaptable to GPU hardware, the authors believe that this computation method is an efficient substitute to full time-resolved reconstructions.

### 3.4.2. Comparing windows and frequency-based reconstruction

One of the question that could arise now is why not to apply the Fourier transform to DTOFs signals and perform the frequency based reconstruction described at Equation 3.3. To answer this question, reconstructions were also performed using Fourier transform datatypes for an inclusion 3.5 cm deep with the same optical values in the previous section.

On the left figure, complex numbers obtained from the Fourier transform were used for the reconstruction. In Figure 3.18 (e) the frequency magnitude obtained at first detector is given. The used frequencies started from 0 MHz up to a maximum value which ranged from 300 MHz to 2 GHz, the frequencies were equally spaced in steps of 100 MHz.

The results show that using frequencies up to 1 GHz (see Figures 3.18 b) yields the best reconstruction in terms of absorption quantification and localization; this agrees with the results of [239] that suggest that useful information can be found at GHz order. Nevertheless, no improvement was obtained by using frequencies up to 2 GHz (see Figure 3.18 a). When only frequencies up to 500 MHz are used (see Figure 3.18 c), the absorption quantification is slightly underestimated, but most of the information is still there, as reported in [109]. If only frequencies up to 300 MHz are included (see Figure 3.18 d) in the reconstruction algorithm, the absorption quantification is highly underestimated compared to previous reconstructions. If only the magnitude of the Fourier transform is used (Figure 3.18 f) the inclusion localization and absorption are highly underestimated; the phase term was not used since it resulted in huge artifacts.

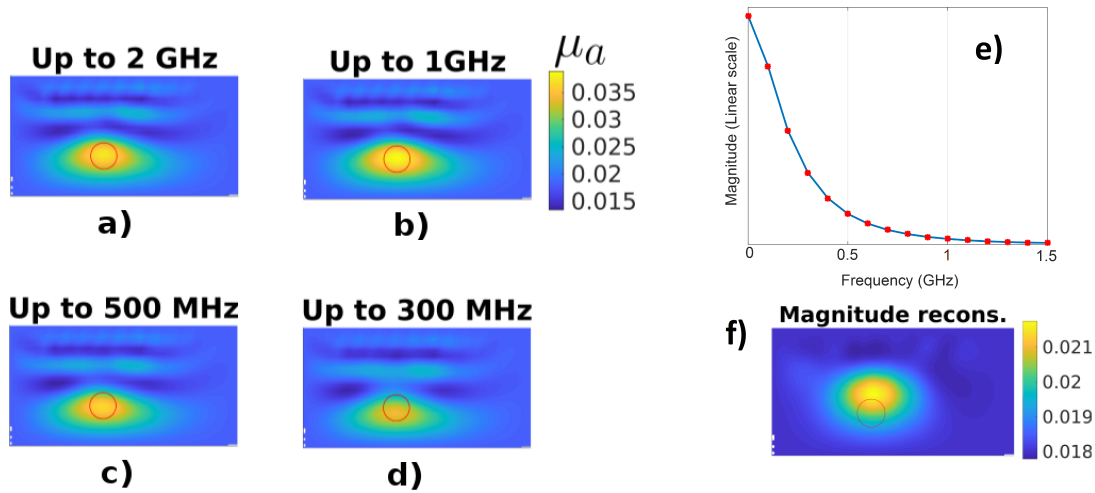


Figure 3.18: (a-d) Absorption reconstructions for 3.5 cm deep inclusion. Maximum used frequencies were 2 GHz, 1GHz, 500 MHz and 300 MHz. (e) Magnitude of the signal at frequency space. (f) Reconstructions using only the magnitude of the Fourier transform. The same optical properties as in previous sections were used.

The equivalence of frequency and windows based results can be understood by realizing that in both cases, the complex numbers associated to each frequency are fitted. In the case of Fourier transform datatype, this is done directly, but for temporal windows, this is done by windowing the frequency domain. Nevertheless, there are some important differences. First, as was shown in Subsection 3.2.4, by applying the Fourier transform to a noisy signal the noise becomes correlated. This does not occur when not overlapping time windows such as Tukey windows are used. Second, when reconstruction is done directly in frequency domain, the linear system to solve is complex while in the temporal windows case it is real. Third, temporal window-based reconstruction is more intuitive to understand, since it is associated to photon arrival times. Therefore, in some cases, a weight matrix could be built in order to weight some windows over other ones, depending on how noisy they are. Last, in the same fashion as the paper published by Hasnain et al. [95], a system that directly measures a gated DTOF (e.g. using rectangular gates) instead of the full curve could be faster than nowadays time-resolved systems. Moreover, those gated DTOFs could be used in the developed reconstruction process since they are equivalent to Tukey windows. Such a system could be potentially faster than state-of-the-art systems without compromising reconstruction quality.

### 3.5. Conclusions

In this chapter, a new method has been proposed for computing temporal windows. The technique consists in computing the temporal windows in the frequency domain instead of in time domain. The main advantage is that better temporal selectivity and non-correlated windows can be used, without compromising computational efficiency considerably. To illustrate the method, Tukey and Gaussian windows were used. The obtained results in a numerical phantom are evidence that these new class of windows have better photon time-arrival selectivity and do not correlate the noise. As a result, the new developed method

could potentially improve the localization and absorption quantification of inclusions as deep as 3.5 centimeters in a reflectance geometry. These results have been demonstrated to be theoretically equivalent, under certain conditions, to frequency-based reconstruction, when frequencies up to GHz order are used.

In the next chapter, I will approach diffuse optical tomography from regularization perspective. Regularization is a critical aspect when solving ill-posed inverse problems. It makes the results stable by applying some assumptions in the reconstruction problem. In this case, I will analyse the performance of total variation regularization applied to diffuse optical tomography problems when the reconstruction mesh is irregular.





# Total variation regularization for diffuse optical tomography

**J**ACQUES SALOMON HADAMARD was an important French mathematician who did an outstanding analysis of partial differential equations. In his work, he defined *well-posed* mathematical problems as problems with a unique solution that changes continuously with initial conditions [91]. He believed that mathematical problems arising from physical processes had to fulfill these requirements. During the first half of the 20th century, Hadamard was proven to be wrong since many problems from fields such as quantum scattering theory, geophysics or astronomy [87, 206] were not well-posed. Nowadays, problems that are not well-posed problems in terms of Hadamard properties are known as *ill-posed* problems [92]. These problems have multiple solutions and are highly sensitive to initial condition perturbations.

For these reasons, in order to solve ill-posed problems, uniqueness has to be forced by making some assumptions. This technique is known as *regularization* and there exist several well-known assumptions, such as to promote smoothness or sparsity of the solution, to force piecewise constant solutions and so on. The suitability of each regularization technique depends on a specific problem and the validity of the assumption.

In this chapter, I will explain how to perform total variation regularization for Diffuse Optical Tomography (DOT) inverse problems on irregular meshes. Most of the content of this chapter is the result of a collaboration I did with PhD student Wenqi Lu and her thesis director PhD Iain Styles from the University of Birmingham. As a result from this collaboration, we published a paper in Biomedical Optics Express journal [135]. This chapter is mainly based on the work described in that paper.

## 4.1. Regularization for DOT inverse problem

The main characteristic of ill-posed problems is that they are very sensitive to perturbations in the initial conditions. Due to this reason, solutions to ill-posed problems are usually regularized by adding some smoothing requirement. In the case of DOT the linear system to solve is known to be an ill-posed problem since the sensitivity matrix is ill-conditioned.

DOT linear system can be approached as an optimization problem by adding a regularizing term as follows:

$$\min_{\mathbf{u}} \|\mathbf{A}\mathbf{u} - \mathbf{b}\|_2^2 + \lambda \mathcal{R}(\hat{\mathbf{u}} + \mathbf{u}), \quad (4.1)$$

where  $\mathbf{A}$  is the sensitivity matrix,  $\mathbf{b}$  is the difference between measurements and simulations,  $\hat{\mathbf{u}} \in \mathbb{R}^N$  is the optical property value defined in a mesh with  $N$  nodes,  $\mathbf{u} \in \mathbb{R}^N$  represents the change of the optical property,  $\mathcal{R}(\hat{\mathbf{u}} + \mathbf{u})$  is the regularization term and  $\lambda$  is the regularization weight.

In DOT, Tikhonov regularization [102] is commonly used due to its implementation easiness but it smears sharp edges of solutions [245]. Sparsity solutions have also been used although its validity to image human tissue has been brought into question [201, 15, 116].  $L_p$  regularization has also been used to promote sparsity [179, 162] but its nonconvexity makes the problem harder to optimize.

Apart from previous techniques, total variation (TV) is one of the most used regularization techniques for image reconstruction [163]. However, its performance on anatomical complex domains has not been analysed in the context of DOT. Total variation regularization minimizes the  $l_1$  norm of the gradient, which promotes piecewise constant solutions. The gradient operator is well-defined in continuous media. However, for discretized domains the gradient operator has to be redefined. A full discussion can be read at [42] for structured meshes, that is, meshes with regular connectivity where neighborhood relationships are well defined. For not structured meshes, as shown at Figure 4.1 (Left), the gradient operator has to be redefined in a different way.

## 4.2. The gradient operator in unstructured meshes: two different approaches

In this section, I describe the two different approaches we proposed to define the gradient operator for unstructured meshes: the finite element and graph-based approaches [45]. For each approach, I will develop the isotropic and anisotropic version of TV regularization.

### 4.2.1. Finite element representation

As was seen in Subsection 2.4.2, in the finite element representation, the geometry is discretized using triangular and tetrahedral elements for two and three-dimensional geometries, respectively. Those elements are composed by a set of vertices and edges and do not overlap with each other. For example, in Figure 4.1 (left) the two-dimensional circumference was discretized using irregular triangles. As shown by the highlighted red triangle, each triangular element is composed by three edges and three vertices.

The finite element representation assumes that infinite-dimensional spaces can be approximated by  $n$ -dimensional subspaces, see Subsection 2.4.2. That is, an infinite-dimensional function can be approximated by the continuous and piecewise-polynomial function  $U(x, y) : \Omega \rightarrow \mathbb{R}$

$$U(x, y) = \sum_{i=1}^N u_i \phi_i(x, y), \quad (4.2)$$

where  $\Omega$  is the computational domain,  $N$  is the number of nodes in the mesh,  $u_i$  represents the optical property value at node  $i$  and  $\phi_i$  is the linear basis function associated to node  $i$ .

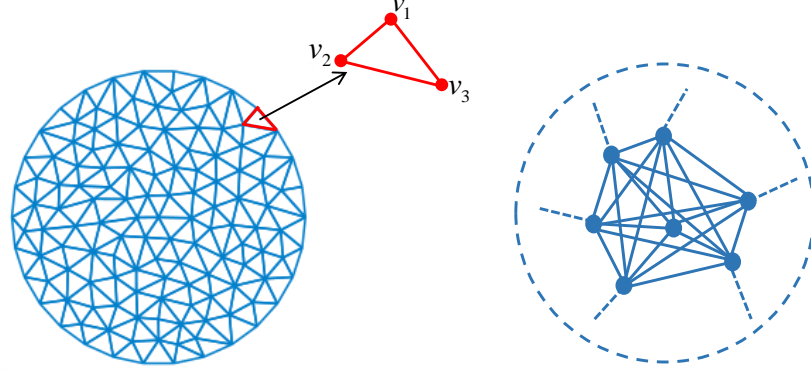


Figure 4.1: (Left) A two-dimensional geometry modelled using finite elements. (Right) Graph-based representation. (Figure taken from [135]).

Using the definition of the  $U(x, y)$  function, the continuous TV form can be represented in discrete form as

$$\int_{\Omega} \left| \frac{\partial U}{\partial x} \right| + \left| \frac{\partial U}{\partial y} \right| dx dy = \|\mathbf{D}_x \mathbf{u}\|_1 + \|\mathbf{D}_y \mathbf{u}\|_1, \quad (\text{anisotropic})$$

$$\int_{\Omega} \sqrt{\left( \frac{\partial U}{\partial x} \right)^2 + \left( \frac{\partial U}{\partial y} \right)^2} dx dy = \sum_{i=1}^M \sqrt{|\mathbf{D}_x \mathbf{u}_i|^2 + |\mathbf{D}_y \mathbf{u}_i|^2}, \quad (\text{isotropic})$$

(4.3)

where  $M$  is the number of elements in the mesh,  $\mathbf{D}_x$  and  $\mathbf{D}_y \in \mathbb{R}^{M \times N}$  are linear operators that map a vector  $\mathbf{u} \in \mathbb{R}^N$  into its partial derivatives. The first equation is the anisotropic version since it promotes directions along the  $x$  and  $y$  axes. The second equation is the isotropic version because it does not favor any particular direction, see Figure 4.2.

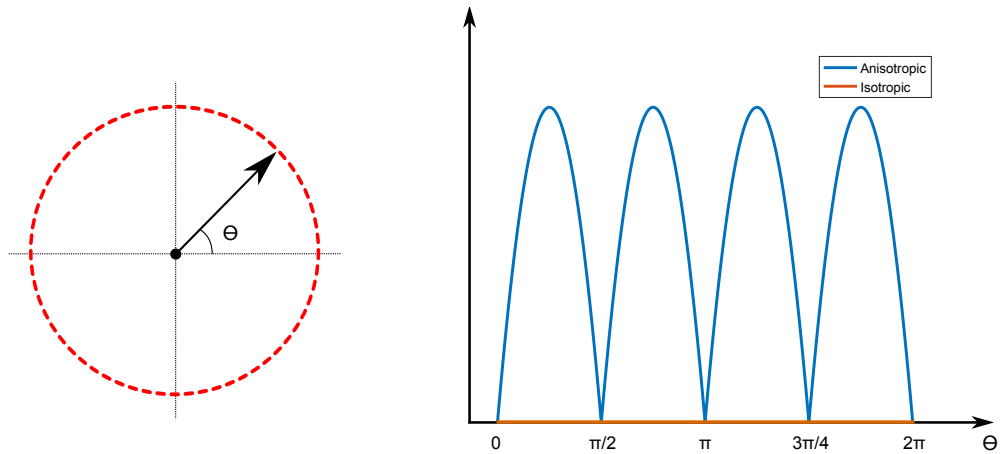


Figure 4.2: Comparison of anisotropic and isotropic norms in a unit circle. The anisotropic norm ( $|x| + |y|$ ) has four minima corresponding to  $0, \pi/2, \pi$  and  $3\pi/4$  angles. The isotropic norm ( $\sqrt{x^2 + y^2}$ ) has minima everywhere since it is constant.

A partial derivative matrix  $\mathbf{D}_x$  or  $\mathbf{D}_y$  for a triangular irregular mesh can be built using an approach similar to the one used to build finite element matrices. For simplicity, I will assume a two-dimensional mesh, although the extension to three dimensions is straightforward.

Given an arbitrary triangle  $\tau$  with vertices at

$$\begin{cases} \mathbf{v}_1 = (x_1, y_1) \\ \mathbf{v}_2 = (x_2, y_2) \\ \mathbf{v}_3 = (x_3, y_3) \end{cases} \quad (4.4)$$

there exists three linear basis functions  $\phi_i$ :

$$\begin{cases} \phi_1(x, y) = a_1x + b_1y + c_1, \\ \phi_2(x, y) = a_2x + b_2y + c_2, \\ \phi_3(x, y) = a_3x + b_3y + c_3, \end{cases} \quad (4.5)$$

where  $a_1 = (y_2 - y_3)/(2A_T)$ ,  $b_1 = (x_3 - x_2)/(2A_T)$ ,  $c_1 = (x_2y_3 - x_3y_2)/(2A_T)$ ,  $a_2 = (y_3 - y_1)/(2A_T)$ ,  $b_2 = (x_1 - x_3)/(2A_T)$ ,  $c_2 = (x_3y_1 - x_1y_3)/(2A_T)$ ,  $a_3 = (y_1 - y_2)/(2A_T)$ ,  $b_3 = (x_2 - x_1)/(2A_T)$  and  $c_3 = (x_1y_2 - x_2y_1)/(2A_T)$ .  $A_T$  denotes the triangular area of  $T$ , which is computed as  $A_T = |x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)|/2$ . The point  $(x, y)$  is assumed to be inside the triangle  $T$ .

Moreover, taking into account that

$$\begin{aligned} \int_{\Omega} \left| \frac{\partial U}{\partial x} \right| dx dy &= \sum_{i=1}^M \int_{T_i} \left| \frac{\partial U}{\partial x} \right| dx dy = \sum_{i=1}^M \int_{T_i} \left| \sum_{j=1}^N u_j \frac{\partial \phi_j}{\partial x} \right| dx dy \\ &= \sum_{i=1}^M A_{T_i} |a_{i,1}u_{i,1} + a_{i,2}u_{i,2} + a_{i,3}u_{i,3}| \\ &= \sum_{i=1}^M |(\mathbf{D}_x \mathbf{u})_i| = \|\mathbf{D}_x \mathbf{u}\|_1, \end{aligned} \quad (4.6)$$

the derivative matrix  $\mathbf{D}_x$  can be built with the following pseudocode:

- **Starting point:** Fill the matrix  $\mathbf{D}_x$  with all zeros.
- For each element  $i$  in the mesh:
  1. Compute the coefficients  $a_1$ ,  $a_2$  and  $a_3$ .
  2. Multiply each coefficient by the area of the element  $A_{T_i}$ .
  3. Go to  $i$ -th row of  $\mathbf{D}_x$  and fill the columns corresponding to the three vertices of the element.
- **Ending point:** After iterating over all elements the partial derivative matrix has been completed.

It works similarly for  $y$  dimension and its extension to three-dimensions is straightforward. The resulting derivative matrices are sparse.

In the next section, I will introduce graph representation of gradient operator and derive the isotropic and anisotropic TV regularization forms.

### 4.2.2. Graph-based representation

An unstructured domain  $\Omega$  can be discretized by a weighted graph  $G = (V, E, w)$ , see Figure 4.1 (right). The graph  $G$  contains a finite set of vertices  $V = \{V_k\}_{k=1}^N$  and a finite set of weighted edges  $E \in V \times V$ . In this work, I assume that  $G$  is an undirected simple graph, that is, there are not multiple edges between two vertices. I also define  $(i, j) \in E$  as the edge of  $E$  that connects the vertices  $i$  and  $j$ .

The graph based discrete differential operators are defined based on nonlocal methods [81]. The nonlocal gradient operator applied to element  $i$  of vector  $\mathbf{u}$  is

$$\nabla_W u_i = (u_j - u_i) \sqrt{W_{ij}} : V \rightarrow \mathbb{R}^N. \quad (4.7)$$

For a vertex  $i \in V$ , the  $\nabla_W u_i$  is a vector with length  $N$ . The weight  $W_{ij} : V \times V \rightarrow \mathbb{R}^+$  represents a weighted link between nodes  $i$  and  $j$ . The matrix  $W$  is symmetric and non-negative.

In this work, the weight between nodes was chosen to be the inverse of the Euclidean distance between each pair of nodes (if  $i = j \rightarrow W_{ij} = 0$ ) because the further the nodes, the lower the influence they have on the gradient. The main difference between finite elements and graph representation is that in the former the gradient has two directions in 2D ( $x$  and  $y$ ) and three in 3D ( $x$ ,  $y$  and  $z$ ), whereas the latter has as many dimensions as edges connected to the node.

The anisotropic graph TV regularization is

$$\sum_{i=1}^N \sum_{j=1}^N \left| (u_j - u_i) \sqrt{W_{ij}} \right|, \quad (4.8)$$

and the isotropic version

$$\sum_{i=1}^N \sqrt{\sum_{j=1}^N (u_j - u_i)^2 W_{ij}}. \quad (4.9)$$

Graph representation allows to use the fully connected version of the graph, that is, that each node is connected with the rest of the nodes. Nevertheless, computationally, this approach is very time-consuming, since the number of edges grows in the order of  $N(N-1)/2$ . In the literature, spectral graph theory [18, 145] and nearest neighbour [67, 28] have been applied to limit the number of nodes used. In this work, it was decided to use the same connectivity between vertices as the mesh used in finite element representation. Using that configuration the graph is sparsified and only neighbour nodes are taken into consideration. Therefore, regularization terms described in Equations 4.8 and 4.9 are converted to

$$\sum_{i=1}^N \sum_{j \in \mathcal{N}_i} \left| (u_j - u_i) \sqrt{W_{ij}} \right|, \quad (4.10)$$

and

$$\sum_{i=1}^N \sqrt{\sum_{j \in \mathcal{N}_i} (u_j - u_i)^2 W_{ij}}, \quad (4.11)$$

where  $\mathcal{N}_i = \{j \in V : (i, j) \in E\}$ .

After presenting the two different approaches to tackle the total variation problem for unstructured meshes, in the next section I will pose their respective optimization problems.

### 4.3. Minimization of TV-regularized DOT inverse problems

TV-regularized inverse problems are not straightforward to minimize since the TV regularization term is not differentiable everywhere. Moreover, standard techniques to minimize  $L_1$  norm problems such as FISTA [17] are not applicable due to the existence of the gradient operator. In the paper, we proposed an algorithm based on Alternating Direction Method of Multipliers (ADMM) algorithm. This algorithm has been widely studied in the last ten years and has been applied into different problems. It consists on splitting the optimization problem in two dual variables and minimize them independently and alternating between each variable, see our paper [135]. Nevertheless, in this chapter, I will describe a different algorithm based on majorize-minimize technique, which approximates the optimization problem with successive quadratic functions. Before starting my collaboration with Wenqi, I have already developed majorize-minimize algorithms, because they can be applied to a larger set of problems without making big changes in the code (later I will show how to adapt the algorithm for sparsity problems). Note that the results are not dependent on the optimization algorithm but only on the used gradient representation.

#### 4.3.1. The optimization problem

The regularization term  $\mathcal{R}(\hat{\mathbf{u}} + \mathbf{u})$  of Equation 4.1 has four different forms: the isotropic and anisotropic version of finite element and graph representation, see Table 4.1. In the next section, I describe the application of majorize-minimize algorithm for TV regularization problems using finite element representation. However, the formulas given can be easily applied to TV forms using graph representation.

| Name   | Formulation  |
|--------|--|
| A-FETV | $\mathbf{u}^* = \arg \min_{\mathbf{u}} \{ \ \mathbf{A}\mathbf{u} - \mathbf{b}\ _2^2 + \lambda \ \mathbf{D}_x(\hat{\mathbf{u}} + \mathbf{u})\ _1 + \lambda \ \mathbf{D}_y(\hat{\mathbf{u}} + \mathbf{u})\ _1 \}$                              |
| I-FETV | $\mathbf{u}^* = \arg \min_{\mathbf{u}} \left\{ \ \mathbf{A}\mathbf{u} - \mathbf{b}\ _2^2 + \lambda \sum_{i=1}^M \sqrt{ (\mathbf{D}_x(\hat{\mathbf{u}} + \mathbf{u}))_i ^2 +  (\mathbf{D}_y(\hat{\mathbf{u}} + \mathbf{u}))_i ^2} \right\}$   |
| A-GTV  | $\mathbf{u}^* = \arg \min_{\mathbf{u}} \left\{ \ \mathbf{A}\mathbf{u} - \mathbf{b}\ _2^2 + \lambda \sum_{i=1}^N \sum_{j \in \mathcal{N}_i}  (\hat{\mathbf{u}} + \mathbf{u})_j - (\hat{\mathbf{u}} + \mathbf{u})_i  \sqrt{W_{ij}} \right\}$   |
| I-GTV  | $\mathbf{u}^* = \arg \min_{\mathbf{u}} \left\{ \ \mathbf{A}\mathbf{u} - \mathbf{b}\ _2^2 + \lambda \sum_{i=1}^N \sqrt{\sum_{j \in \mathcal{N}_i} ((\hat{\mathbf{u}} + \mathbf{u})_j - (\hat{\mathbf{u}} + \mathbf{u})_i)^2 W_{ij}} \right\}$ |

Table 4.1: TV regularization problems using finite element and graph-based approaches. A-FETV, I-FETV, A-GTV and I-GTV respectively represent anisotropic finite element total variation, isotropic finite element total variation, anisotropic graph total variation and isotropic graph total variation.

### 4.3.2. Majorize-minimize algorithm

The developed algorithm is designed using the conjugate gradient (CG) algorithm to find the solution to the minimization problem and approximating  $L_1$  terms by the majorize-minimize approach. This setup allows to solve TV and sparsity problems without making big code changes.

CG method is a well-known optimization technique. Although its linear version is the most popular it can also be applied to nonlinear cases. A simple sketch of CG algorithm is as follows:

1. Set a initial  $\mathbf{x}_0$ .
2. Calculate  $\mathbf{G}_0 = -\nabla f(\mathbf{x}_0)$ .
3. Calculate  $\sigma$  that minimizes  $f(\mathbf{x}_0 + \sigma \mathbf{G}_0)$ .
4. Update  $\mathbf{x}_i = \mathbf{x}_0 + \sigma \mathbf{G}_0$ .
5. Calculate  $\mathbf{G}_i = -\nabla f(\mathbf{x}_i)$ .
6. Use a conjugate formula to obtain  $\beta_i$  (for example Fletcher-Reeves formula).
7. Update the conjugate gradient  $\mathbf{S}_i = \mathbf{G}_i + \beta_i \mathbf{S}_{i-1}$  where  $(\mathbf{S}_0 = \mathbf{G}_0)$ .
8. Calculate  $\sigma$  that minimizes  $f(\mathbf{x}_i + \sigma \mathbf{S}_i)$ .
9. Update position  $\mathbf{x}_{i+1} = \mathbf{x}_i + \sigma \mathbf{S}_i$ .
10. If convergence is reached then stop; if not go to step 5.

### Anisotropic form using majorize-minimize technique

I will develop the majorize-minimize approach for the anisotropic form of TV. The anisotropic version is equivalent to

$$\varepsilon = \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda \|\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x})\|_1 \quad (4.12)$$

$$= \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda \sum_{i=1} \sqrt{\left( \sum_{j=1} B_{ij}(\hat{u}_j + x_j) \right)^2}, \quad (4.13)$$

where  $\mathbf{A}$  is the sensitivity matrix,  $\hat{\mathbf{u}}$  is a constant vector,  $\mathbf{B}$  is the gradient matrix whose rows 1 to  $M$  are  $\mathbf{G}_x$  and rows  $M + 1$  to  $2M$  are  $\mathbf{G}_y$ , and the second equality arises from the definition of the  $L_1$ -norm. Note that I have denoted  $\mathbf{u} = \mathbf{x}$  to promote clarity in the following equations.

If the vector  $\mathbf{x}$  is perturbed by  $\delta\mathbf{x}$  where  $\delta x_i \ll x_i$  for all  $i$  components, then the function value changes by  $\delta\varepsilon$ ,

$$\varepsilon + \delta\varepsilon = \|\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) - \mathbf{b}\|_2^2 + \lambda \|\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x} + \delta\mathbf{x})\|_1 \quad (4.14)$$

$$= \|\mathbf{A}(\mathbf{x} + \delta\mathbf{x}) - \mathbf{b}\|_2^2 + \lambda \sum_{i=1} \sqrt{\left( \sum_{j=1} B_{ij}(\hat{u}_j + x_j + \delta x_j) \right)^2}. \quad (4.15)$$



Then, using the equality  $|x + \delta x| = |x| + x \delta x |x|^{-1} = |x| + \text{sign}(x)\delta x$  the perturbation of the function can be expressed as

$$\delta\varepsilon = 2\delta\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - 2\delta\mathbf{x}^T \mathbf{A}^T \mathbf{b} + \delta\mathbf{x}^T \mathbf{A}^T \mathbf{A} \delta\mathbf{x} + \lambda \sum_{i=1} \frac{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j)\right) \left(\sum_{j=1} B_{ij} \delta x_j\right)}{\sqrt{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j)\right)^2}}. \quad (4.16)$$

Moreover, dismissing second order terms and writing  $E_{ii} = \sqrt{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j)\right)^2}$  as a diagonal matrix the perturbation is as follows

$$\delta\varepsilon \approx 2\delta\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} - 2\delta\mathbf{x}^T \mathbf{A}^T \mathbf{b} + \lambda \sum_{i=1} \frac{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j)\right) \left(\sum_{j=1} B_{ij} \delta x_j\right)}{E_{ii}} \quad (4.17)$$

$$= 2\mathbf{x}^T \mathbf{A}^T \mathbf{A} \delta\mathbf{x} - 2\mathbf{b}^T \mathbf{A} \delta\mathbf{x} + \lambda \left( (\hat{\mathbf{u}} + \mathbf{x})^T \mathbf{B}^T \mathbf{E}^{-1} \mathbf{B} \delta\mathbf{x} \right) \quad (4.18)$$

$$= \mathbf{d}^T \delta\mathbf{x}, \quad (4.19)$$

where the approximation of the gradient is  $\mathbf{d} = 2\mathbf{A}^T \mathbf{A} \mathbf{x} - 2\mathbf{A}^T \mathbf{b} + \lambda \mathbf{B}^T \mathbf{E}^{-1} \mathbf{B} (\hat{\mathbf{u}} + \mathbf{x})$ .

Once the gradient is known, the next step is to compute the conjugate direction  $\mathbf{s}$  with Fletcher-Reeves or Polak-Ribière formulas. After, the optimal step length in the direction of descent must be estimated. That is, the following function must be minimized,

$$\begin{aligned} f(\sigma) &= \varepsilon(\mathbf{x} + \sigma\mathbf{s}) = \|\mathbf{A}(\mathbf{x} + \sigma\mathbf{s}) - \mathbf{b}\|_2^2 + \lambda \|\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x} + \sigma\mathbf{s})\|_1 \\ &= \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{x} + 2\sigma \mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{s} - 2\mathbf{x}^T \mathbf{A}^T \mathbf{b} + \sigma^2 \mathbf{s}^T \mathbf{A}^T \mathbf{A} \mathbf{s} \\ &\quad - 2\sigma \mathbf{s}^T \mathbf{A}^T \mathbf{b} + \mathbf{b}^T \mathbf{b} + \lambda \|\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x} + \sigma\mathbf{s})\|_1. \end{aligned} \quad (4.20)$$

The last term of the function is not derivable at all points; majorize-minimize method approximates the absolute value by a quadratic function  $H(\mathbf{x}, \mathbf{x}_0)$  (majorizing approximation) with the following properties [108]:

- $H(\mathbf{x}, \mathbf{x}_0) \geq |\mathbf{x}|, \forall \mathbf{x}_0, \mathbf{x} \in \mathbb{R}^n,$
- $H(\mathbf{x}_0, \mathbf{x}_0) = |\mathbf{x}_0|, \forall \mathbf{x}_0 \in \mathbb{R}^n,$

which implies that  $|\mathbf{x}_1| \leq H(\mathbf{x}_1, \mathbf{x}_0) \leq |\mathbf{x}_0|$  where  $\mathbf{x}_1 = \arg \min_{\mathbf{x}} H(\mathbf{x}, \mathbf{x}_0)$ .

In this case, the quadratic function will also have the same derivative value as the absolute function at points  $-\mathbf{x}_0$  and  $\mathbf{x}_0$ . The majorizing approximation for one-dimensional case is

$$|x| \leq \frac{1}{2}|x_0| + \frac{1}{2} \frac{x^2}{|x_0|}. \quad (4.21)$$

In Figure 4.3, I plotted the absolute value function and several majorizing approximations.

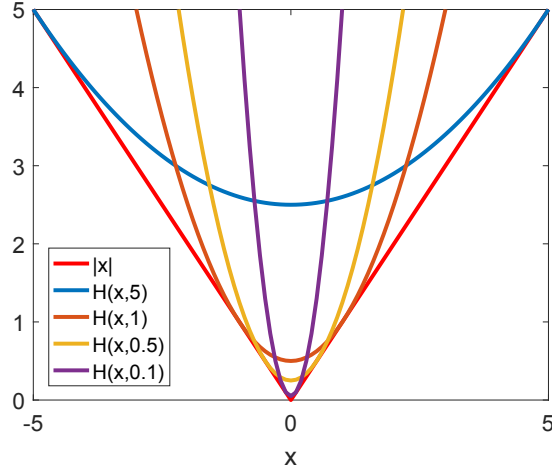


Figure 4.3: Absolute value function and majorizing approximation at different  $|x_0|$  values.

Using the majorizing approximation, the term  $\|\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x} + \sigma\mathbf{s})\|_1$  can be approximated as,

$$\begin{aligned}
\|\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x} + \sigma\mathbf{s})\|_1 &= \sum_{i=1} \sqrt{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j + \sigma s_j)\right)^2} \\
&\leq \frac{1}{2} \sum_{i=1} \sqrt{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j)\right)^2} + \sum_{i=1} \frac{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j + \sigma s_j)\right)^2}{2\sqrt{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j)\right)^2}} \\
&= \sum_{i=1} \sqrt{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j)\right)^2} \\
&\quad + \sum_{i=1} \frac{\sigma^2 \left(\sum_{j=1} B_{ij} s_j\right)^2 + 2\sigma \left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j)\right) \left(\sum_{j=1} B_{ij} s_j\right)}{2\sqrt{\left(\sum_{j=1} B_{ij}(\hat{u}_j + x_j)\right)^2}} \\
&= \|\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x})\|_1 + \sigma(\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x}))^T \mathbf{E}^{-1} \mathbf{B} \mathbf{s} + \frac{\sigma^2}{2} (\mathbf{B} \mathbf{s})^T \mathbf{E}^{-1} \mathbf{B} \mathbf{s}.
\end{aligned} \tag{4.22}$$

So,

$$\begin{aligned}
f(\sigma) &= \|\mathbf{A}(\mathbf{x} + \sigma\mathbf{s}) - \mathbf{b}\|_2^2 + \lambda \|\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x} + \sigma\mathbf{s})\|_1 \\
&\leq \|\mathbf{A}(\mathbf{x} + \sigma\mathbf{s}) - \mathbf{b}\|_2^2 + \lambda \left( \|\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x})\|_1 + \sigma(\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x}))^T \mathbf{E}^{-1} \mathbf{B} \mathbf{s} + \frac{\sigma^2}{2} (\mathbf{B} \mathbf{s})^T \mathbf{E}^{-1} \mathbf{B} \mathbf{s} \right) \\
&= F(\sigma).
\end{aligned} \tag{4.23}$$

Therefore, taking the derivative with respect to  $\sigma$ :

$$\begin{aligned} \frac{\partial F(\sigma)}{\partial \sigma} &= 2\mathbf{x}^T \mathbf{A}^T \mathbf{A} \mathbf{s} - 2\mathbf{s}^T \mathbf{A}^T \mathbf{b} + \lambda(\mathbf{B}(\hat{\mathbf{u}} + \mathbf{x}))^T \mathbf{E}^{-1} \mathbf{B} \mathbf{s} + \sigma (2\mathbf{s}^T \mathbf{A}^T \mathbf{A} \mathbf{s} + \lambda(\mathbf{B} \mathbf{s})^T \mathbf{E}^{-1} \mathbf{B} \mathbf{s}) \\ &= \mathbf{s}^T \mathbf{s} + \sigma (2\mathbf{s}^T \mathbf{A}^T \mathbf{A} \mathbf{s} + \lambda(\mathbf{B} \mathbf{s})^T \mathbf{E}^{-1} \mathbf{B} \mathbf{s}) = 0, \end{aligned} \quad (4.24)$$

and the optimal step length is approximated as

$$\sigma = \frac{-\mathbf{s}^T \mathbf{s}}{2\mathbf{s}^T \mathbf{A}^T \mathbf{A} \mathbf{s} + \lambda(\mathbf{B} \mathbf{s})^T \mathbf{E}^{-1} \mathbf{B} \mathbf{s}}. \quad (4.25)$$

This is the algorithm for the anisotropic case when derivatives with respect to  $x$  and  $y$  are managed as separate entities. In the next section, I will develop the isotropic form.

### Isotropic form using majorize-minimize technique

The isotropic norm for a matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is defined as:

$$\|\mathbf{A}\|_{1,2} = \sum_{i=1}^n \left( \sum_{j=1}^m a_{ij}^2 \right)^{1/2}, \quad (4.26)$$

which is known as the  $L_{2,1}$  matrix norm.

For the TV problem, the norm is defined as

$$\|\mathbf{D}\mathbf{u}\|_{1,2} = \sum_{i=1}^M \left( \left( \sum_{j=1}^N (\mathbf{D}_x)_{i,j} u_j \right)^2 + \left( \sum_{j=1}^N (\mathbf{D}_y)_{i,j} u_j \right)^2 \right)^{1/2}. \quad (4.27)$$

where  $\mathbf{D}_x, \mathbf{D}_y \in \mathbb{R}^{M \times N}$  are the gradient matrices for  $x$  and  $y$  components respectively.

Then, isotropic version of the TV problem is equivalent to

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2 + \lambda \|\mathbf{D}(\hat{\mathbf{u}} + \mathbf{x})\|_{1,2}. \quad (4.28)$$

Performing similar mathematical operations as for the anisotropic form, the gradient for the isotropic norm is:

$$\mathbf{d}^T = 2(\mathbf{A}\mathbf{x} - \mathbf{b})^T \mathbf{A} + \lambda \mathbf{D}_x^T \mathbf{E}^{-1} \mathbf{D}_x (\hat{\mathbf{u}} + \mathbf{x}) + \lambda \mathbf{D}_y^T \mathbf{E}^{-1} \mathbf{D}_y (\hat{\mathbf{u}} + \mathbf{x}) \quad (4.29)$$

and the step is

$$\sigma = \frac{-\mathbf{s}^T \mathbf{s}}{2\mathbf{s}^T \mathbf{A}^T \mathbf{A} \mathbf{s} + \lambda (\mathbf{D}_x \mathbf{s})^T \mathbf{E}^{-1} \mathbf{D}_x \mathbf{s} + \lambda (\mathbf{D}_y \mathbf{s})^T \mathbf{E}^{-1} \mathbf{D}_y \mathbf{s}}, \quad (4.30)$$

where  $\mathbf{s}$  is the conjugate direction and

$$E_{ii} = \sqrt{\left( \sum_{j=1}^N (D_x)_{i,j} \hat{u}_j \right)^2 + \left( \sum_{j=1}^N (D_y)_{i,j} \hat{u}_j \right)^2}. \quad (4.31)$$

### Sparsity form

It should be noted that for the anisotropic norm when  $\mathbf{B} = \mathbf{I}$ , the sparsity form arises. Therefore, the gradient and optimal steps length are:

$$\min_{\delta\mu_a} \|\mathbf{Ax} - \mathbf{b}\|_2^2 + \lambda \|\hat{\mathbf{u}} + \mathbf{x}\|_1. \quad (4.32)$$

Then, the direction of descent and optimal step length are

$$\mathbf{d}^T = 2(\mathbf{Ax} - \mathbf{b})^T \mathbf{A} + \lambda(\hat{\mathbf{u}} + \mathbf{x})^T \mathbf{E}^{-1}, \quad (4.33)$$

$$\sigma = \frac{-\mathbf{s}^T \mathbf{s}}{2\mathbf{s}^T \mathbf{A}^T \mathbf{A} \mathbf{s} + \lambda \mathbf{s}^T \mathbf{E}^{-1} \mathbf{s}}, \quad (4.34)$$

where

$$E_{ii} = \sqrt{(\hat{u}_i + x_i)^2}. \quad (4.35)$$

As can be seen, the main advantage of this technique is that different optimization can be solved without having to adapt significantly the code. Moreover, it is also possible to perform optimization adding total variation and sparsity terms at the same time.

## 4.4. Experiments

In this section, I make a qualitative and quantitative comparison of the finite element and graph-based approaches. First, I describe four different evaluation metrics to quantify the quality of the reconstructed images. Then, to answer the scientific objectives, I show the numerical simulations we did in a circular 2D domain and 3D head model, see Figures 4.4 (a) and (b) respectively. Finally, I describe the experiments performed in a cylindrical phantom, see Figure 4.4 (c). In 2D, the mesh was composed by triangular elements (see Figure 4.4 (d)) and in 3D by tetrahedral elements (see Figure 4.4 (e)).

The simulations were performed using the NIRFAST package [54] in Matlab R2017a. They were all based on single-wavelength continuous-wave technology and were corrupted by Gaussian noise; ten repetitions were performed to represent the variability of the results.

### 4.4.1. Quantitative evaluation metrics

There were used four different evaluation metrics: the localization error, average contrast, peak signal-to-noise ratio and relative recovered volume. These metrics allow to measure the reconstructions quality from different points of view.

Localization error is defined as Euclidean distance between the center of mass of simulated activation region  $\mathbf{x}_s$  and the center of mass of the recovered activation region  $\mathbf{x}_r$ . If the reconstructed image is identical to the ground truth image, the localization error is zero. The recovered activation region is determined by selecting the recovered changes that are greater than 60% of the maximum recovered changes,

$$\text{Localization error} = \|\mathbf{x}_s - \mathbf{x}_r\|_2. \quad (4.36)$$

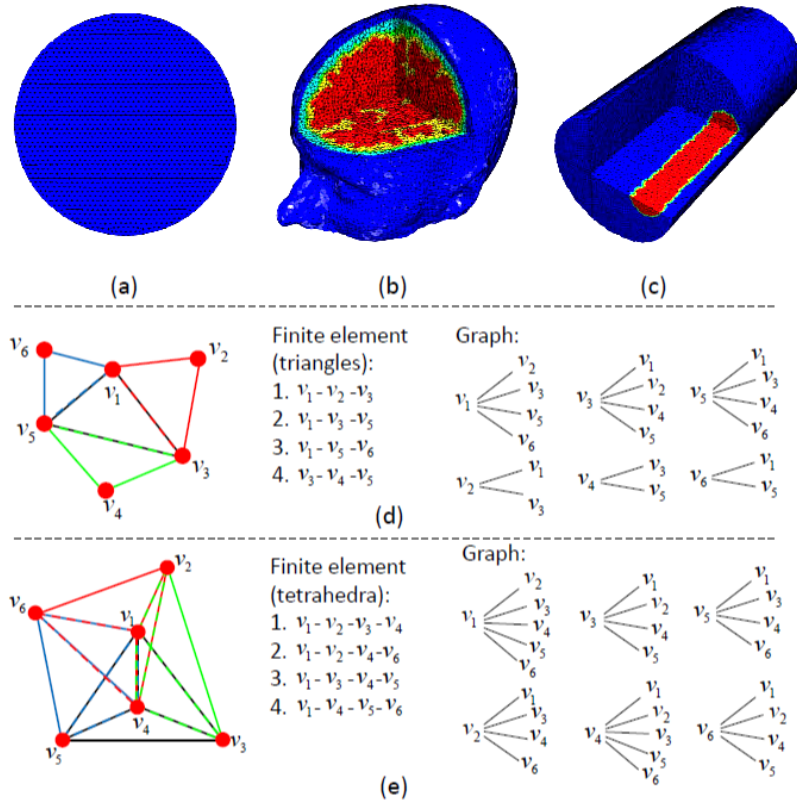


Figure 4.4: (a)-(c): Discrete computational domain of the three geometries. (d) Example of mesh composition for a 2D geometry in the finite element and graph representation respectively. (e): Example of mesh composition for a 3D geometry in the finite element and graph representation respectively. (Figure taken from [135]).

The average contrast evaluation metric is based on the mean value of the region of interest:

$$\text{Average contrast} = \frac{\sum_{i \in N_r} \mu_i}{|N_r| \hat{\mu}}, \quad (4.37)$$

where  $\mu_i$  denotes the recovered optical property at node  $i$ ,  $N_r$  represents the set of nodes at the region of interest,  $|N_r|$  is the cardinality of the set and  $\hat{\mu}$  is the ground truth values of the optical property at the region of interest. If the reconstructed image is identical to the ground truth image, the average contrast value is one.

The third evaluation metric is the peak signal to noise ratio (PSNR). PSNR is defined as follows:

$$\text{PSNR} = 10 \cdot \log_{10} \left( (\max(\mu))^2 / \text{MSE} \right), \quad (4.38)$$

where MSE is the mean squared error between the recovered and ground truth images, that is,  $\text{MSE} = \sum_{i=1}^N (\mu_i - \hat{\mu}_i)^2 / N$ . Usually, the larger the value of PSNR the less difference between the recovered and ground truth. Nevertheless, as this metric can poorly perform for high value outliers it should be taken with a pinch of salt.

The last evaluation metric is the relative recovered volume ( $V_{RRV}$ ) which is defined as

$$V_{RRV} = V_r / V_s \times 100\%, \quad (4.39)$$

where  $V_r$  and  $V_s$  are the volume of the activation region for recovered and simulated cases. The volume of the recovered region,  $V_r$ , is computed by thresholding the recovered changes based on 60% of the maximum and computing the volume of those elements. If the reconstructed image is identical to the ground truth image, the relative recovered volume is one.

#### 4.4.2. Two-dimensional simulations

The objective of the two-dimensional simulations is to test the accuracy of finite element and graph approaches. Moreover, since the number of nodes and elements in two-dimensional space is significantly lower than three-dimensional domains it can also be tested for high resolution meshes.

First, the TV anisotropic version of both approaches using two different mesh resolutions was used. After, identical tests were performed using the isotropic TV version.

##### Anisotropic TV regularization

Anisotropic TV assumes that the shape of the reconstructed object is orientated with the coordinate axis. That is, it promotes horizontal and vertical structures, since oblique objects increase the TV measure [85]. This assumption may not be accurate for DOT problems since usually inclusions are not blocks. Nevertheless, it will be interesting to see the differences between the two approaches and how they manage the same problem.

The two-dimensional circular geometry (Figure 4.4 (a)) had a radius of 43 mm. At position  $(-10 \text{ mm}, 10 \text{ mm})$  an inclusion of radius 10 mm was placed. Around the boundary sixteen source-detector fibers separated equidistantly were placed. The light sources were assumed to have a constant intensity over time, that is, they were continuous-wave sources. When one of the fibers is switch-on as a source the rest are assumed to be detectors leading to 240 total boundary data points per wavelength. Sources were placed  $1/\mu'_s$  in the domain due to their spherically isotropic behaviour in Diffusion Approximation equation.

Two meshes with different resolutions were used in the simulation. The coarser mesh had 1785 nodes and 3418 linear triangle elements with the average element size  $0.6977 \text{ mm}^2$  (Figure 4.5 (a)) while the finer one had 5133 nodes and 10013 elements with the average element size  $0.5801 \text{ mm}^2$  (Figure 4.5 (d)). The background absorption coefficient  $\mu_a$  is set as  $0.1 \text{ cm}^{-1}$  and for the anomaly is  $0.3 \text{ cm}^{-1}$  (Figure 4.5 (b) and (e)). Reduced scattering coefficient  $\mu'_s$  remained constant as  $10 \text{ cm}^{-1}$ . The measurements were corrupted with normally distributed Gaussian noise ranging from 0% to 3% at 1% steps of percentage noise.

Reconstructed absorption values are shown in Figure 4.5 (c) and (f). A-FETV reconstructs the inclusion by preserving the blocky shape aligned with the coordinates. This is the expected behaviour when anisotropic TV is used. Nevertheless, A-GTV reconstructs the inclusion by preserving the original shape, it has no bias to any direction. This behaviour is logical since graph representation does not use the concept of direction but it adapts the gradient depending on the edges. That is the reason why reconstructed object shape changes with the resolution of the mesh. Although, the outcome is better than for A-FETV this is not how an anisotropic norm should behave. Therefore, if blocky reconstructions are

preferred finite element representation should be used since it preserves the anisotropy of the norm.

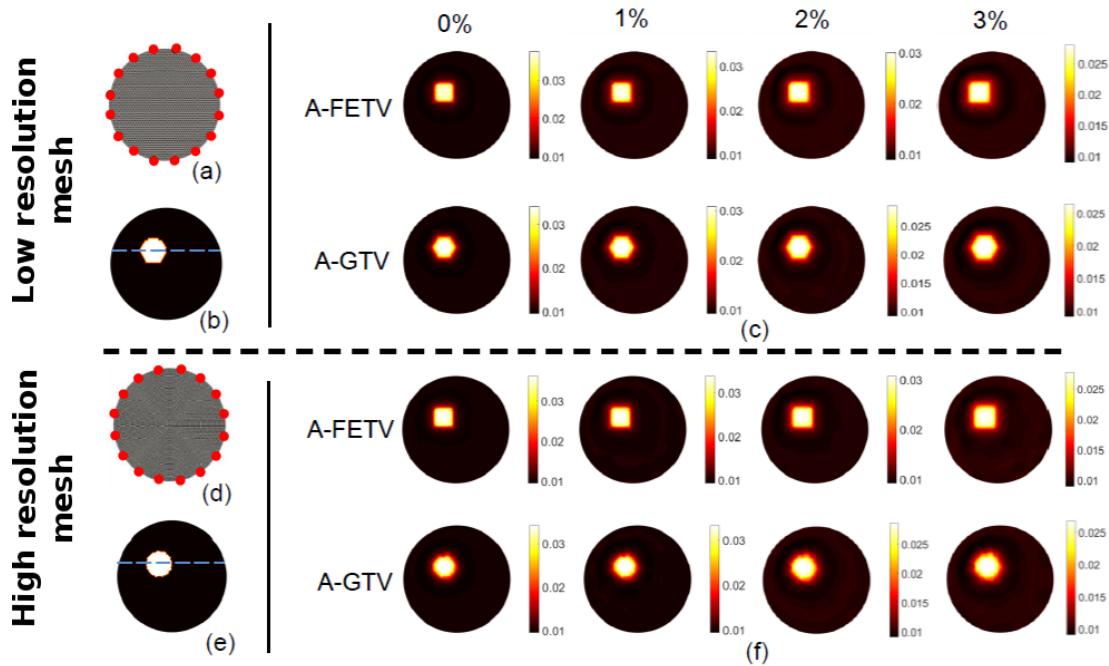


Figure 4.5: (a): Two-dimensional mesh with low spatial resolution. Red points in the boundary indicate the sources and detectors positions. (b): Original absorption distribution for low resolution mesh. (c) Absorption reconstructions with low spatial resolution mesh using anisotropic TV. First and second row use finite element and graph approaches respectively. Columns represent measurements with different levels of noise corruption. (d): High spatial resolution two-dimensional mesh. (e): Original absorption distribution for high resolution mesh. (f): Absorption reconstructions with high spatial resolution mesh using anisotropic TV. (Figure taken from [135]).

### Isotropic TV regularization

The same setup as for previous case was used. In this case, isotropic TV regularization was performed and Tikhonov regularization was included as baseline. In Figure 4.6 the two dimensional case for isotropic TV norm is shown. Tikhonov regularization tends to over-smooth the solutions and create some oscillatory artifacts in the background. When the reconstruction mesh has a low resolution (Figure 4.6 (c)) the difference between I-FETV and I-GTV approaches is very small. However, when the mesh has a larger resolution the I-FETV method performs better than I-GTV, see Figure 4.6 (f). Similar findings are obtained from cross sections results shown at Figure 4.7. Tikhonov regularization reconstructs a peak solution with over-smoothed boundaries. However, TV techniques reconstruct piecewise constant images, which converge to the solution as the resolution of the mesh is larger. It must be noted that finite element based solution (blue curve) adapts much better to the corners than the graph based solution (red line) for high resolution mesh. In Figure 4.8 the quantitative metrics results are shown; the areas represent the 25% to 75% value among the ten repetitions. It is seen that the performance of I-FETV improves with an increase of mesh resolution: by 25% in localization error, 26% in average contrast and 11% in PSNR,

meanwhile the performance of I-GTV is independent of the mesh resolution. These results show that I-FETV is affected considerably by the mesh resolution, the better the resolution the better the reconstructions. However, I-GTV shows more independence to the mesh resolution. As will be shown later, for 3D cases this factor will be of great advantage when I-GTV is used.

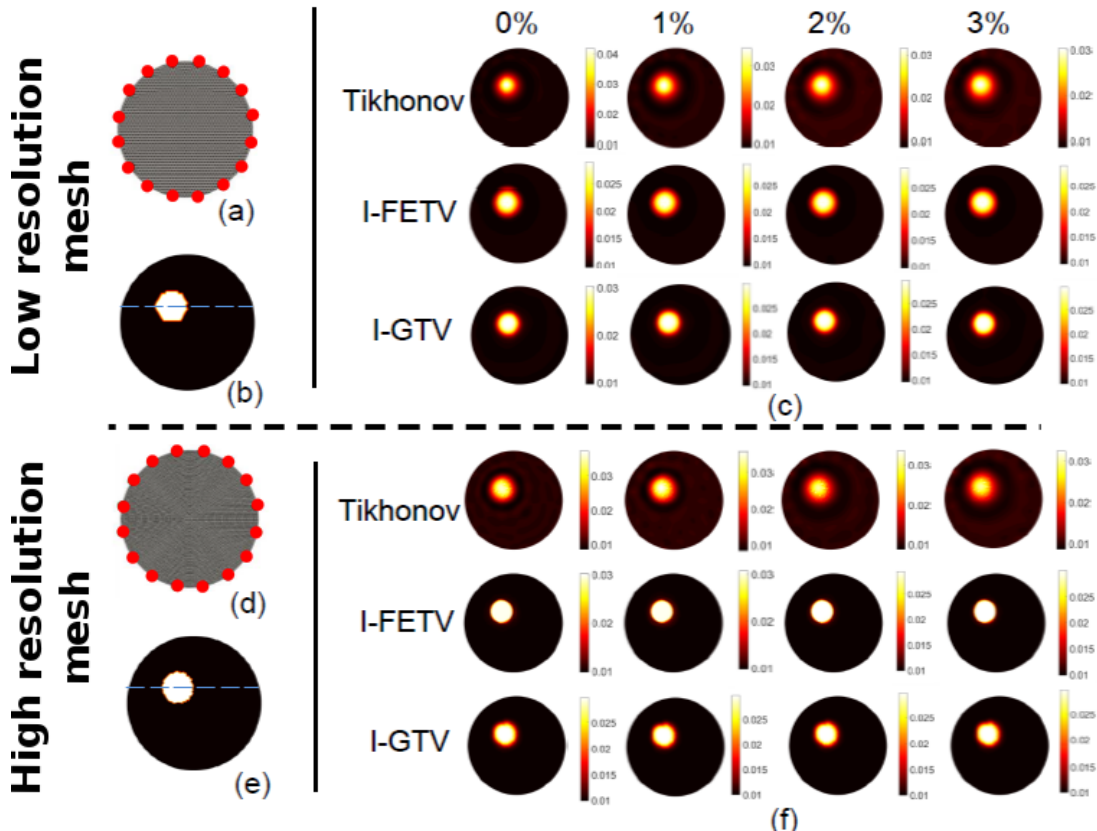


Figure 4.6: Same setup as Figure 4.5 but using isotropic TV regularization. Tikhonov regularization was also added in the first row as reference. (Figure taken from [135]).

#### 4.4.3. Three dimensional head simulations

The TV isotropic regularization techniques were also tested in a realistic 3D head model. This model represents a physically realistic head and its resolution is not very large since the number of nodes increases drastically the computational time of simulations. The model was built from T1-weighted MPRAGE scans [64] in combination with Statistical Parametric Mapping (SPM) software [166] to perform parametric segmentation of head tissues (scalp, skull, cerebrospinal fluid (CSF), gray matter, white matter). SPM result was processed with NIRFAST [111] to create a volumetric FEM mesh based on tetrahedra. The mesh consisted of 50721 nodes and 287547 tetrahedral elements, with an average element size of  $9.2676 \text{ mm}^3$ . Each node was labelled with a head tissue type. Absorption coefficients of each head tissue were obtained from an in-vivo study [211, 48, 21] using 750 nm wavelength laser, see Table 4.2.



|                              | Scalp | Skull | CSF  | Grey Matter | White Matter |
|------------------------------|-------|-------|------|-------------|--------------|
| $\mu_a$ (cm <sup>-1</sup> )  | 0.17  | 0.12  | 0.04 | 0.18        | 0.17         |
| $\mu'_s$ (cm <sup>-1</sup> ) | 7.4   | 9.4   | 3    | 8.4         | 11.9         |

Table 4.2: Head tissue optical properties. Data taken from in-vivo studies [211, 48, 21]

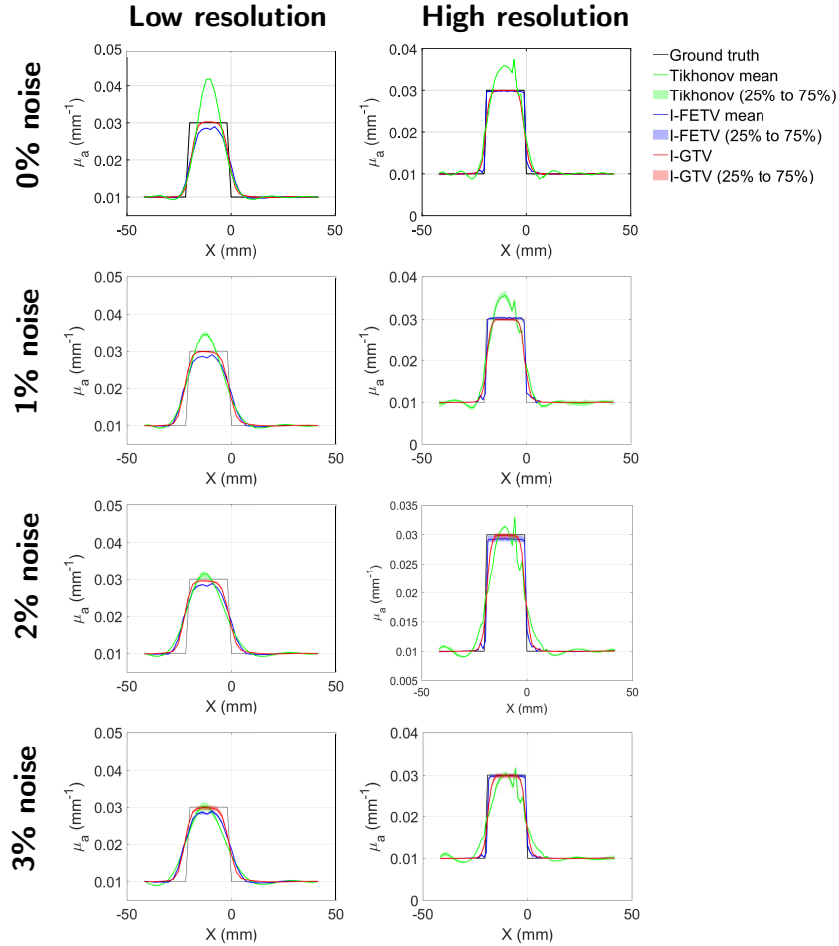


Figure 4.7: One dimensional cross section of absorption distributions recovered at Figure 4.6 (cross-section was performed at the blue line shown at Figure 4.6 (b,e)). First and second column correspond to low and high resolution meshes respectively. Meanwhile, first to fourth row correspond to 0%, 1%, 2% and 3% levels of Gaussian noise. The shadow areas represent the 25% to 75% value among the ten repetitions. (Figure taken from [135]).

A high-density imaging array with 158 sources and 166 detectors (see Figure 4.10 Left) was placed over the whole head. The source-detector separation distances ranged from 1.3 to 4.8 cm, see Figure 4.9. In this work, 3478 differential measurements per wavelength were used to image absorption changes. Both frontal and back anomalies were placed in the brain, each of them with 15 mm radius. The absorption coefficient of the anomalies were calculated by taking into account that in traumatic brain injury cases the tissue oxygen saturation (StO<sub>2</sub>) is normally between 50% and 75% [107]. In this case it was assumed that StO<sub>2</sub> was 55%. In order to provide realistic simulations in accordance with in vivo measurements,

0.12%, 0.15%, 0.41% and 1.42% Gaussian random noise was added to first (13mm), second (30mm), third (40mm) and fourth (48mm) nearest neighbor (NN) measurements [56], see Figure 4.9. Reconstructed absorption images are displayed in Figure 4.10.

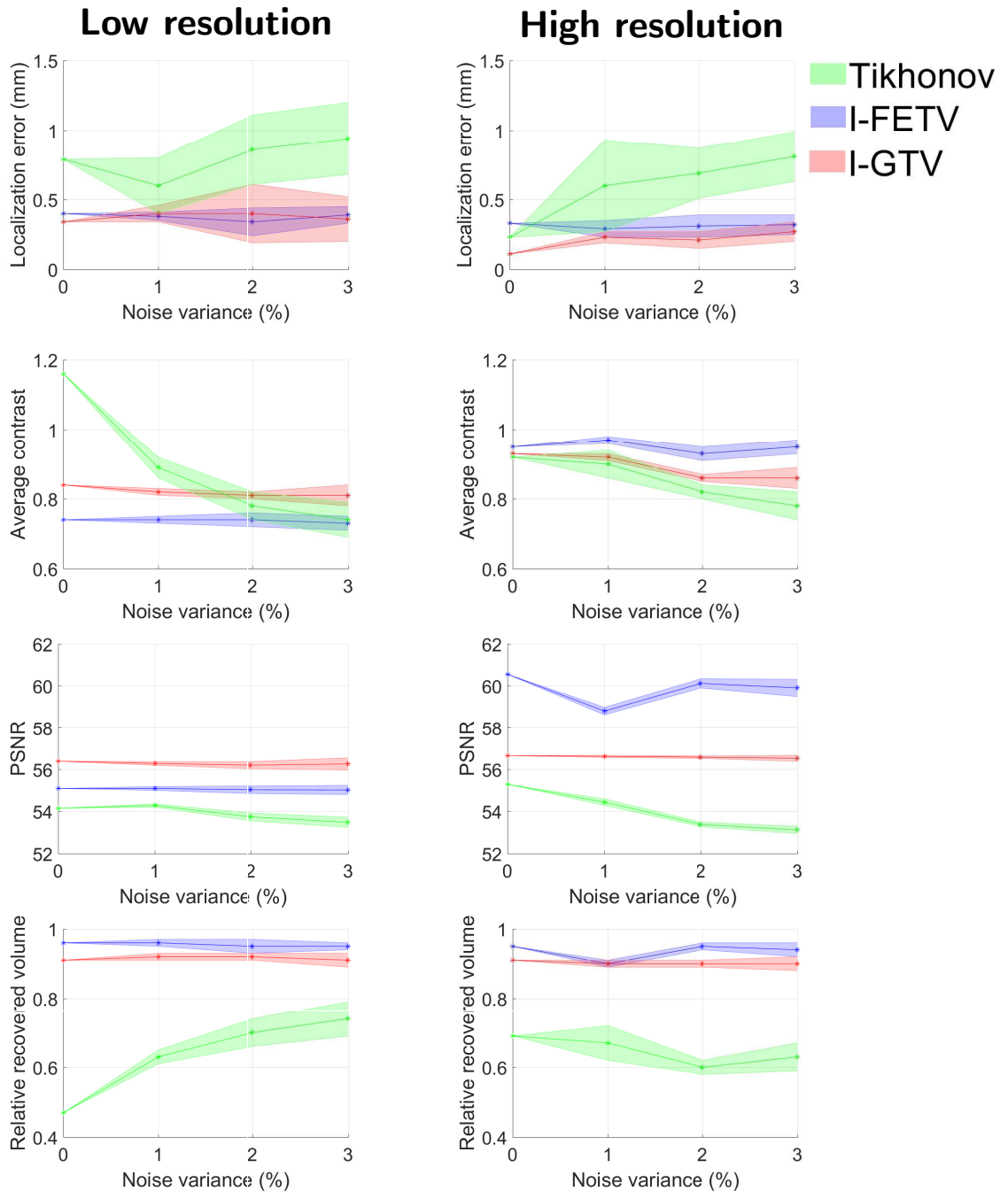


Figure 4.8: Evaluation metrics comparing the performance of Tikhonov (green), isotropic finite element TV (blue) and isotropic graph TV (red) regularization techniques. Left column corresponds to the reconstructions with low resolution mesh in Figure 4.6. Right column corresponds to high resolution mesh in Figure 4.6. The shadow areas represent 25% to 75% value among ten repetitions. (Figure taken from [135]).

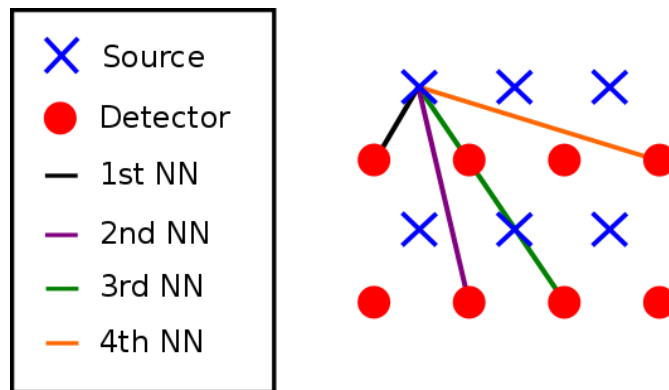


Figure 4.9: Source detector setup. For each nearest neighbor distance a Gaussian random noise of 0.12%, 0.15%, 0.41% and 1.42% was added to first (13mm), second (30mm), third (40mm) and fourth (48mm) distance respectively.

The Tikhonov regularization leads to many artifacts close to the boundary (see the slices of the head shown at Figures 4.11 and 4.12). Moreover, anomalies are not well localized between the boundaries because of the over-smoothing that Tikhonov imposes. I-FETV and I-GTV do not suffer from these surface artifacts and tightly localize the anomalies. Although, from reconstructed images there is no obvious difference between I-FETV and I-GTV approaches, the evaluation metrics displayed at Figure 4.13 show that I-GTV achieves the lowest localization error, largest peak signal-to-noise ratio and a better average contrast. These results confirm the conclusions we obtained from two-dimensional experiment: I-GTV performs slightly better than I-FETV when the resolution of the mesh is low.

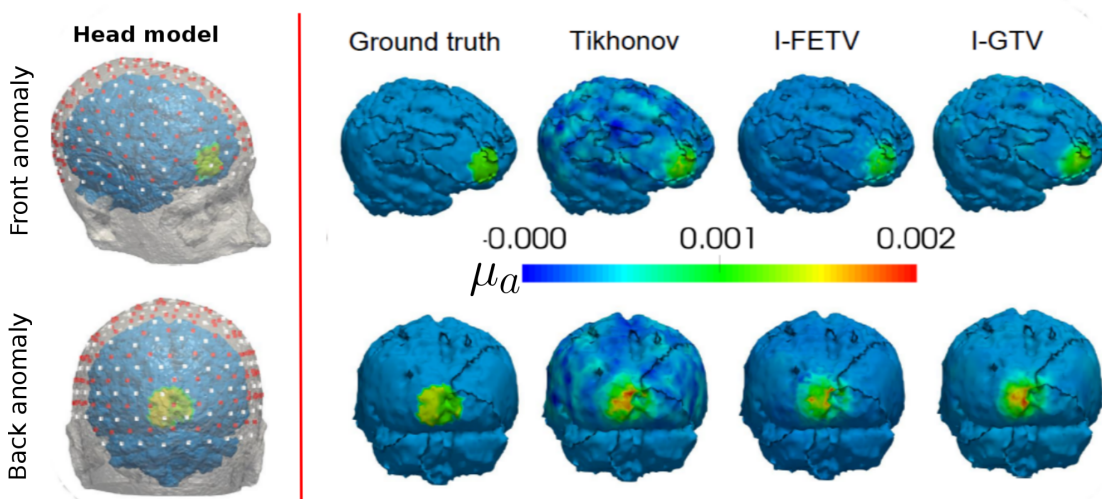


Figure 4.10: In the first column it is shown the source and detector distribution located in the head (158 sources, red dots and 166 detectors, white dots). The positions and absorption values of simulated anomalies are shown in second column. Reconstructed absorption images with Tikhonov, I-FETV and I-GTV are shown in third, fourth and fifth columns respectively. (Figure taken from [135]).

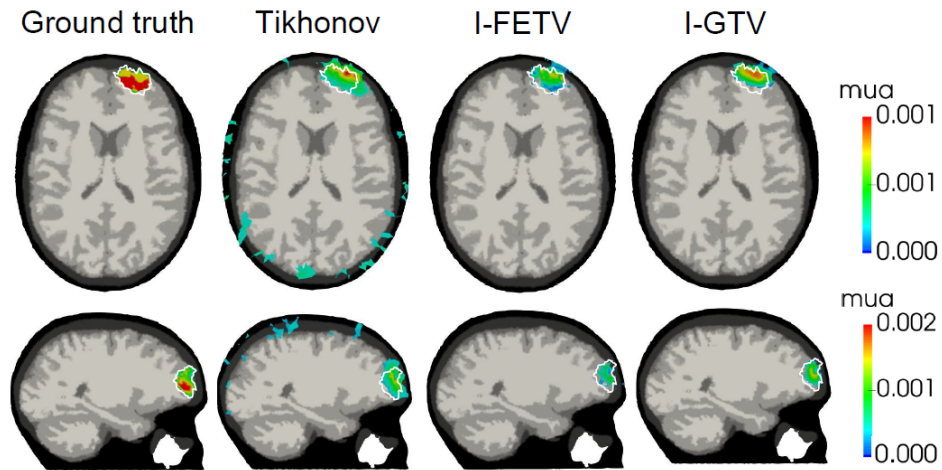


Figure 4.11: Two-dimensional slices for the anomaly located at the forehead. The true anomaly location was highlighted with a white boundary. The absorption and scattering of the anomaly at different brain layers is given at Table 4.2. Note that colorbar values were rounded up to first three decimal digits. (Figure taken from [135]).

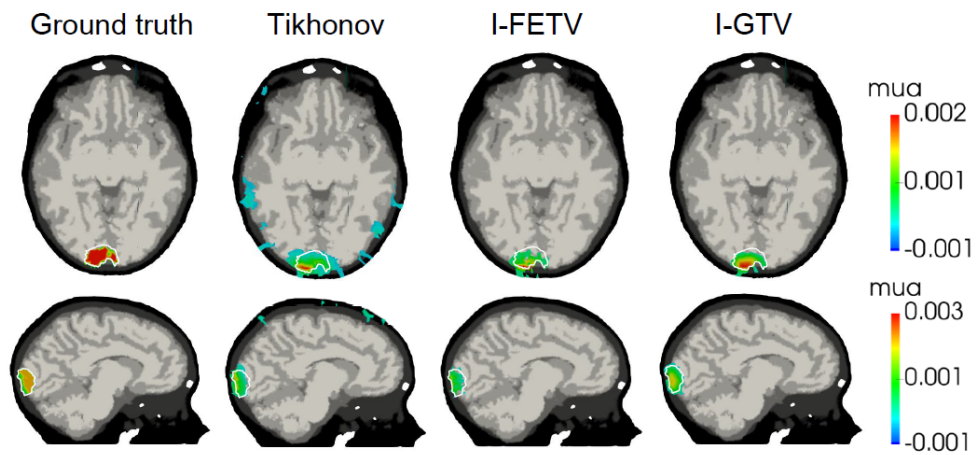


Figure 4.12: Two-dimensional slices for the anomaly located at the back-head. The true anomaly location was highlighted with a white boundary. (Figure taken from [135]).

#### 4.4.4. Solid phantoms

The TV regularization methods were also tested on experimental data. This data was measured from a solid plastic cylindrical phantom using a non-contact CW-DOT system. The length of the phantom is 50 mm and it has a radius of 12.3 mm. 35 sources and 99 detectors are positioned on the underside and top of the phantom respectively (Figure 4.14 (a)). The absorbing dye within the phantom was treated as a chromophore that has unit concentration in the bulk of the phantom. A cylindrical rod was placed at the depth of 5 mm to simulate the heterogeneous inclusion of the phantom (Figure 4.14 (a)). The inclusion rod has a length of 50 mm and a radius of 3 mm and provides a 2:1 contrast in dye concentration compared to background. The system uses five different wavelengths (650 nm, 710 nm, 730 nm, 830 nm

and 930 nm). The reconstruction mesh consists of 9082 nodes and 48099 linear tetrahedral elements with the average tetrahedral elements size  $0.4218 \text{ mm}^3$ . The ground truth data and reconstructed images using different methods are shown in Figure 4.14 (c). These results and the metrics at Table 4.3 confirm the conclusions from previous simulations. In this case only the central part of the rod was reconstructed since the sources and detectors had very low sensitivity to absorption changes outside the central part. Tikhonov regularization shows artifacts close to the sources and detectors and tends to create less uniform solutions outside the inclusion. Meanwhile, TV based regularization shows a more uniform reconstruction with fewer artifacts. Specially, I-GTV recovers a very homogeneous solution outside the anomaly boundaries.

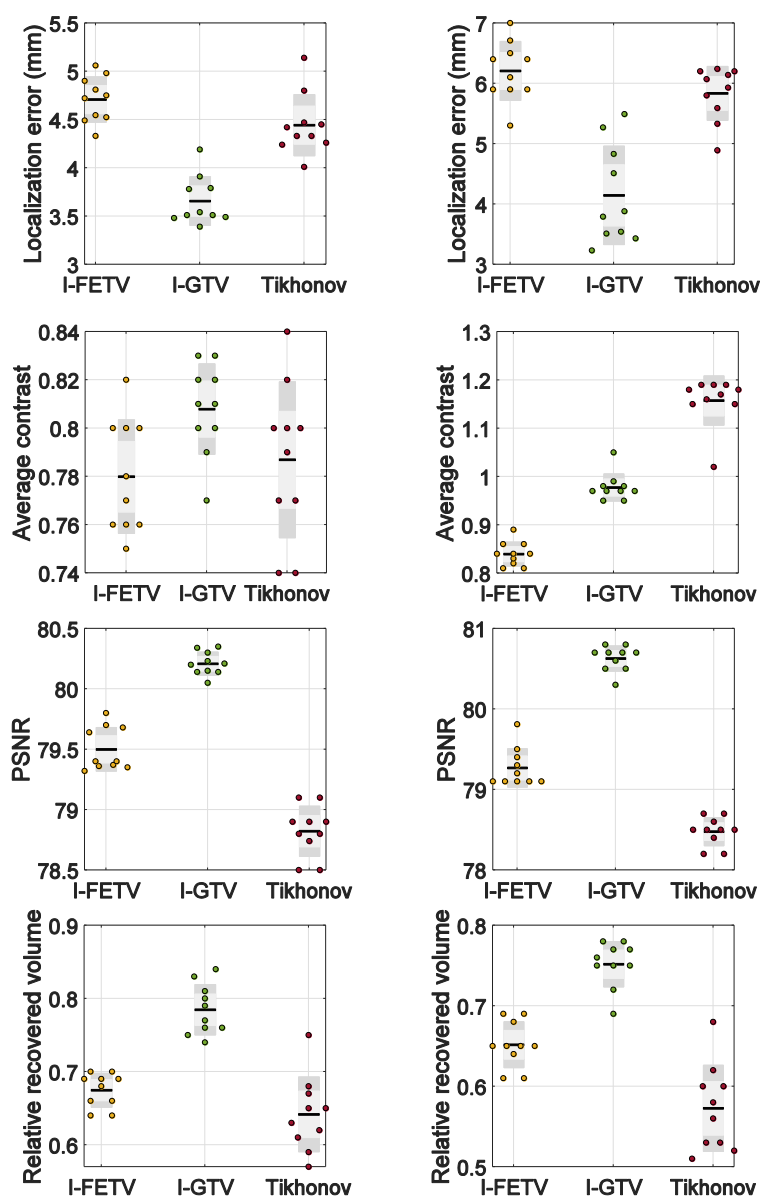


Figure 4.13: Evaluation metrics comparing the performance of different methods on a 3D head model. Left and right column represents the results obtained for forehead and back-head anomaly respectively. (Figure taken from [135]).

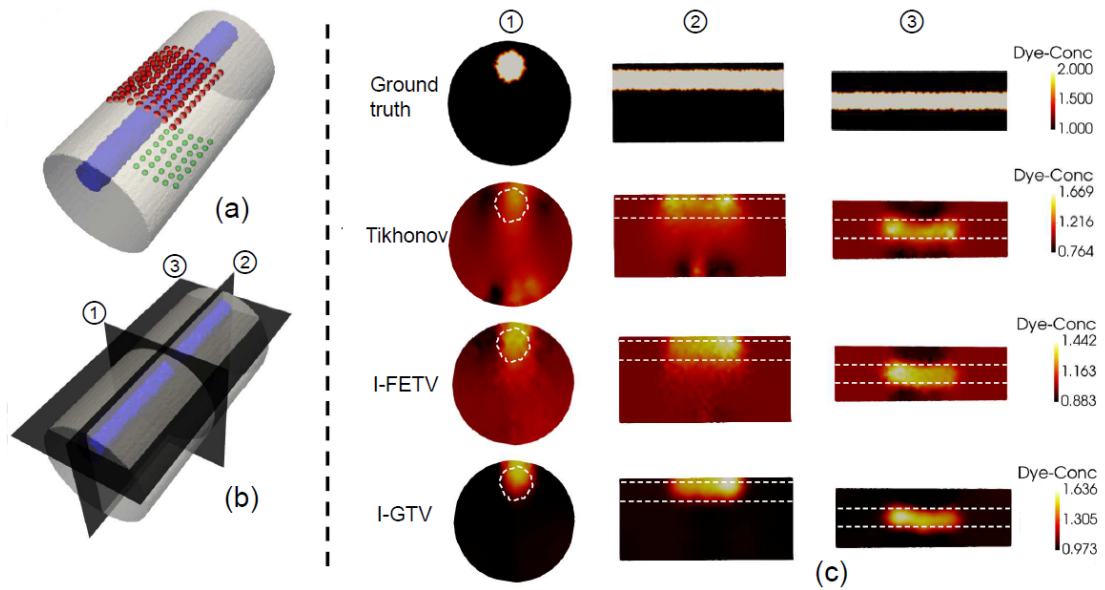


Figure 4.14: (a) Cylindrical geometry with sources and detectors. (b) The different view plans. (c) From top to bottom: ground truth, Tikhonov, I-FETV and I-GTV. The columns represent each view planes. (Figure taken from [135]).

|          | Localization error [mm] | Average contrast [-] | PSNR [-] | Relative recovered volume [%] |
|----------|-------------------------|----------------------|----------|-------------------------------|
| Tikhonov | 2.90                    | 0.74                 | 13.74    | 40                            |
| I-FETV   | 2.81                    | 0.69                 | 14.77    | 48                            |
| I-GTV    | 3.16                    | 0.79                 | 16.71    | 46                            |

Table 4.3: Evaluation metrics values for the cylindrical phantom.

## 4.5. Conclusions

In this chapter, I introduced TV regularization for diffuse optical tomography problems. As part of this collaboration, we proposed two different approaches to represent the gradient in an irregular mesh: the finite element and graph-based approaches. In this work, we also tested the anisotropic and isotropic version of TV. It was shown that the anisotropic TV leads to blocky reconstructions for finite element representation. However, in the graph representation this effect did not have happen since directions are dependant on the mesh edges orientations. The numeric and experimental results for isotropic TV give the same conclusions: (1) Tikhonov regularization tends to oversmooth solution and create oscillatory artifacts, (2) I-GTV creates better constant piecewise solutions than I-FETV for coarse meshes and (3) I-FETV converges faster to ground truth when the mesh resolution is large enough.

In the next chapter, I perform reconstruction from in-vivo measurements obtained in collaboration with Politecnico di Milano. In this experimental work, I performed arm cuff occlusion and motor cortex activation experiments with three different volunteers. As will be shown

later, some regularization techniques will be needed to perform the reconstructions of such experiments.

## Motor cortex activation experiments with a tomographic time-resolved diffuse optical system

**D**URING March–April 2019, I had the opportunity to stay for one month at the Department of Physics from Politecnico di Milano. The goal of my stay at Milan was to perform in-vivo measurements with a recently developed time-resolved diffuse optical system. One of the greatest features of this lab prototype is that the light input is automatically switched between eight fibers in less than 30 ms. Moreover, it also has eight silicon photomultiplier detectors. These characteristics make the system ideal for performing continuous and in-vivo brain tomography of human subjects.

The core idea of the collaboration was to perform diverse in-vivo measurements with Politecnico's device and to apply the developed diffuse optical tomography algorithms. In this chapter, I first describe the time-resolved optical system that was used for in-vivo experiments. Then, I analyse the tomographic capabilities of different source-detector probe geometries. I also explain the data analysis and tomographic reconstruction that were used. After, the instrument and algorithms are validated by performing venous and arterial arm occlusion experiments. Finally, the performed motor cortex activation experiments are described and a detailed data analysis and tomographic reconstructions are given at the end of the chapter.

### 5.1. Experimental setup

The system is based on two wavelengths (670 and 820 nm) pig-tailed laser diode which are driven by the same laser driver (PDL 800, Picoquant GmbH, Germany). Each output is equipped with its collimation stage which hosts also the variable optical attenuator to change the laser power injected in the fiber (100  $\mu\text{m}$  core diameter). For each laser output, the output fiber was hold on a tilting holder (for aligning purposes) and then one plano-convex lens ( $f = 30\text{ mm}$ , diameter 1 inch, Thorlabs GmbH, Germany) was inserted to collimate the beam exiting from the fiber. To have both wavelengths into the same path, a dichroic mirror



at 45 degrees, with cut off wavelength at 805 nm was put. The 820 nm beam was reflected by the dichroic mirror while the 670 nm was transmitted. For this reason, the two output fiber were disposed accordingly and the length of the laser fiber output was chosen to have the two wavelengths one after the other with few nanoseconds difference (first: 820 nm; second: 670 nm). The beam exiting from the dichroic mirror was then focused onto the 100  $\mu\text{m}$  core-input fiber of the  $1 \times 9$  switch using the same plano-convex lens as before to obtain a 1:1 imaging. The  $1 \times 9$  switch (PiezosystemJena GmbH, Germany) fastly switches ( $< 30$  ms) the input on one among the 9 outputs (one of them serving as a stop channel) and it is software controlled. Each of the 8 outputs corresponds to one injection point although only 6 of them were used (three per probe, see Subsection 5.1.1). In the same way, in each probe there were 4 detectors to detect the re-emitted photons. Each detector is a Silicon Photomultiplier (SiPM) with an active area of  $1.3 \times 1.3 \text{ mm}^2$  which can be put directly in contact with the sample, see Figure 5.1 for a schematic view of the optical system. The fibers and detectors were arranged in two separate probes. Details about the detector and its read-out electronics can be find in [70]. The building of the histogram is done by an 8-channel Time-to-Digital (TDC) converter (SC-TDC-1000, Surface Concept GmbH, Germany) featuring an average bin size of 82.2 ps. Other details about the pre-processing of the histogram generated by the TDC can be found in [70, 72].

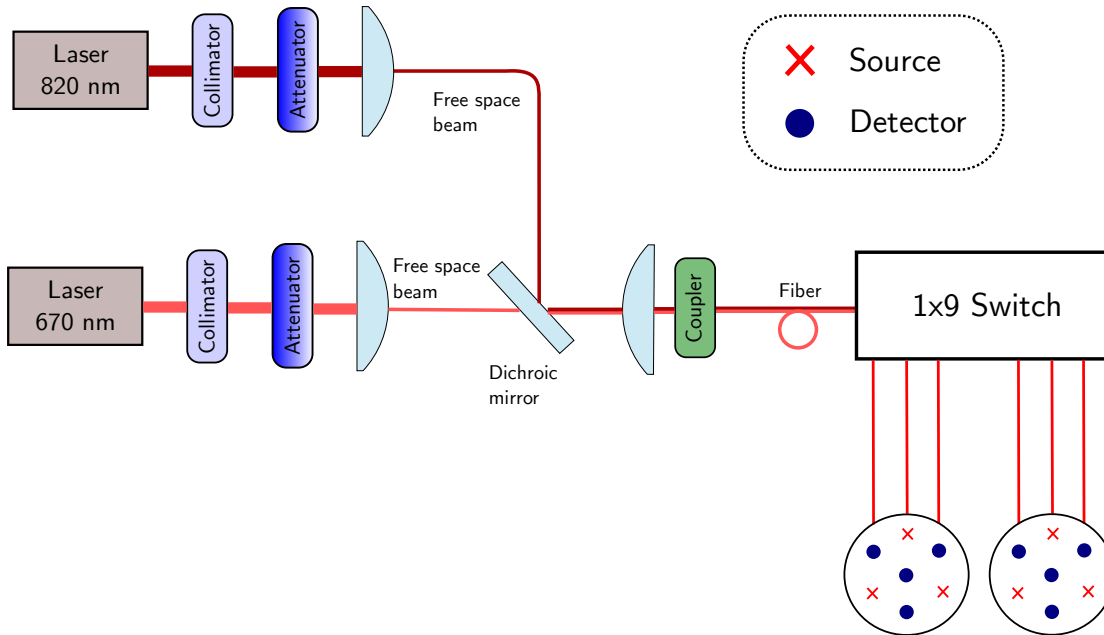


Figure 5.1: The optical system used in arm occlusion and motor cortex activation experiments. It consists of two probes with three fibers and four detectors each.

### 5.1.1. Probe geometry analysis

As was described in the previous section, the optical system has up to eight light fibers and eight SiPM detectors. One of the first challenges to face is which source–detector geometry provides better reconstruction results. To answer this question, I performed simulations using four different geometries. The simulation domain had a size of  $8 \times 8 \times 4 \text{ cm}^3$ . The optical properties of the background were  $\mu_a = 0.1 \text{ cm}^{-1}$  and  $\mu'_s = 10 \text{ cm}^{-1}$ . A spherical inclusion of radius 0.5 cm with optical properties  $\mu_a = 0.3 \text{ cm}^{-1}$  and  $\mu'_s = 10 \text{ cm}^{-1}$  was located at a

depth of 1.5 cm; the goal was to simulate a brain cortex activation. The inclusion was located at four different positions:  $\mathbf{x}_1 = (x, y, z) = (1, 0, 1.5)$ ,  $\mathbf{x}_2 = (0, 0, 1.5)$ ,  $\mathbf{x}_3 = (-1, 0, 1.5)$  and  $\mathbf{x}_4 = (-1, 1, 1.5)$  cm. The tested geometries are shown in Figure 5.2.

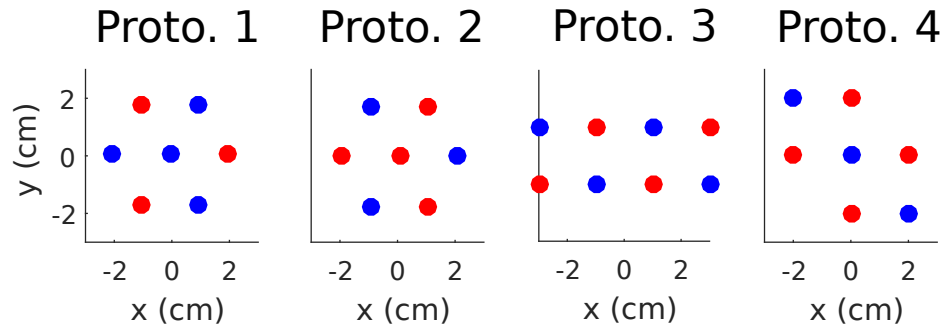


Figure 5.2: Geometry of the four prototypes tested in the simulations. Red dots indicates sources while blue dots are detector positions.

In Figures 5.3, 5.4, 5.5 and 5.6 the inclusion absorption reconstructions are shown at a  $z = 1.5$  cm plane. The green circle indicates the inclusion position and the red/blue dots indicate the source/detector positions at the surface. Reconstructions were performed using Mellin–Laplace moments with  $p = 9$  and up to  $n = 30$  orders. Nodes sensitivity was normalized using the approach described at Subsection 5.2.3. For all probe geometries the minimum distance between source and detector was 2 cm. In the reconstructions, curves obtained from distances greater than 2 cm were not used because in the experimental measurements the obtained curves were usually very noisy.

In Figure 5.3, the results for first prototype are shown. It can be seen that inclusion is well localized for all positions. Nevertheless, when inclusion is located at position  $\mathbf{x}_3$ , the reconstructed inclusion is much bigger than the original and the maximum is not located at the exact position. The reason is that position  $\mathbf{x}_3$  is a “blind spot” of the probe geometry. A blind spot is defined as a location where the probability of many photons travelling from any source to any detector and passing through that area is low. Note that the highest probability zones are the areas in between a source and a detector, e.g. location  $\mathbf{x}_1$ . The blind points of first probe are the positions in between any couple of detectors, such as position of inclusion  $\mathbf{x}_3$ . In Figure 5.7 the blind spots of all geometries are shown. A 0.5 cm thick layer was located in between 1 to 1.5 cm depth. As can be seen, the areas that are not reconstructed are known as blind spots. For prototype 1 and 2, the center point is also a blind spot.

At Figure 5.4 the reconstructions for second prototype are shown. Here the conclusions are similar to previous geometry. At position  $\mathbf{x}_3$  there is a blind spot. Results for this geometry are a little bit worse because for inclusion at position  $\mathbf{x}_2$  the location is slightly shifted.

The results of third geometry are shown at Figure 5.5. It can be seen that a blind spot is located at position  $\mathbf{x}_2$ . The reconstructed  $\mathbf{x}_2$  inclusion is much broader than the original one. Moreover, when inclusion is located under a detector the reconstruction are also quite broad, see reconstruction for inclusion located at  $\mathbf{x}_4$ . At Figure 5.6 the results for fourth geometry are shown. As expected, the blind spot is located at position  $\mathbf{x}_4$ .

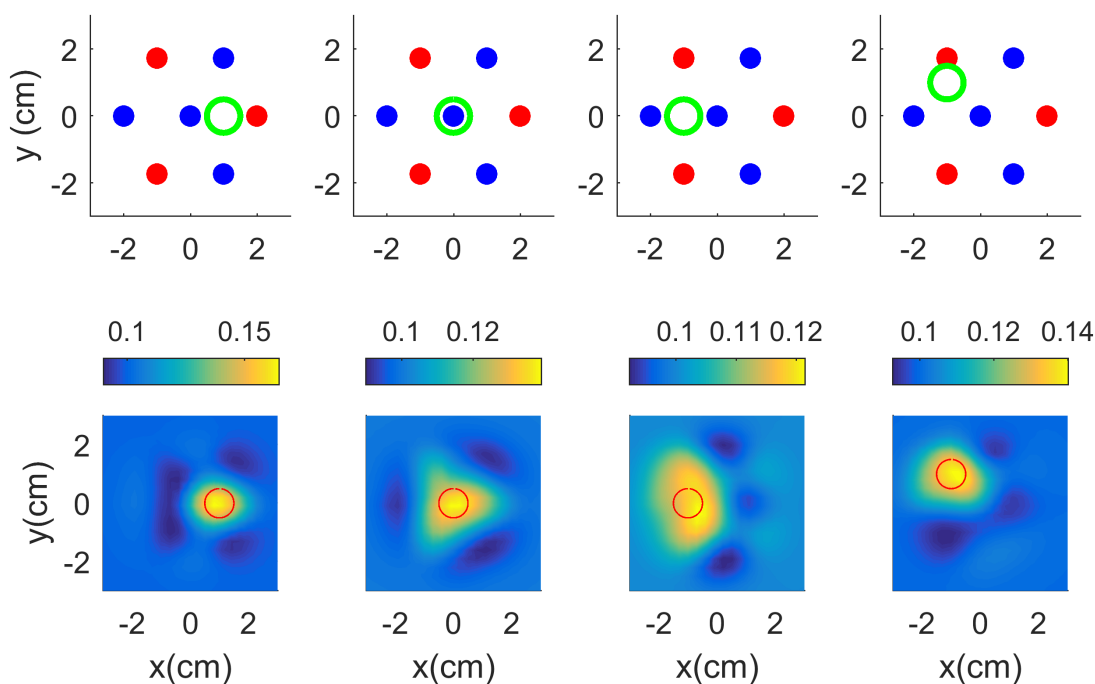


Figure 5.3: Reconstructions for first probe prototype. At first row, the red/blue dots indicate source/detector positions. The green circle indicates the position of the spherical inclusion at  $z = 1.5$  cm plane. At second row, the absorption reconstructions at  $z = 1.5$  cm plane are shown. The red circle indicates the original inclusion position. Colorbars quantify the reconstructed absorption. For visualization purposes, the size of the  $x - y$  plane was cut to  $3 \times 3$  cm size.

All the geometries analysed have blind spots. Those cannot be avoided unless several source–detector distances are used. In this case, this was not possible due to the low signal-to-noise ratio (SNR) at far distances. Between the geometries shown in this work, I decided to select the first geometry prototype (see Figure 5.8) because of the following reasons:

- It has only three blind spots, e.g. the positions in between the center detector and periphery detectors. Although, for large inclusions a blind spot can appear below the central detector. This is also true for the second and third geometry.
- Four detectors are used and not three such as for second geometry. This is important because in some cases a detector could not be fully attached to the surface tissue or count rate is not enough. Therefore, the more detectors there are, the less significant it will be to lose information from one of the detectors.
- The geometry covers a more localized area in comparison with third probe. The motor cortex activation areas are well localized in the brain, therefore the first geometry covers better these areas. Nevertheless, the third probe covers a rectangular area which does not fit so naturally on motor cortex positions. Moreover, the geometry of fourth prototype is less intuitive and it is not clear how to orientate it in the head.
- For first probe prototype, three detectors corresponded to each source, that is, the number of source–detector pairs are equally distributed for each source. However, for the rest of prototypes, this is not fulfilled.

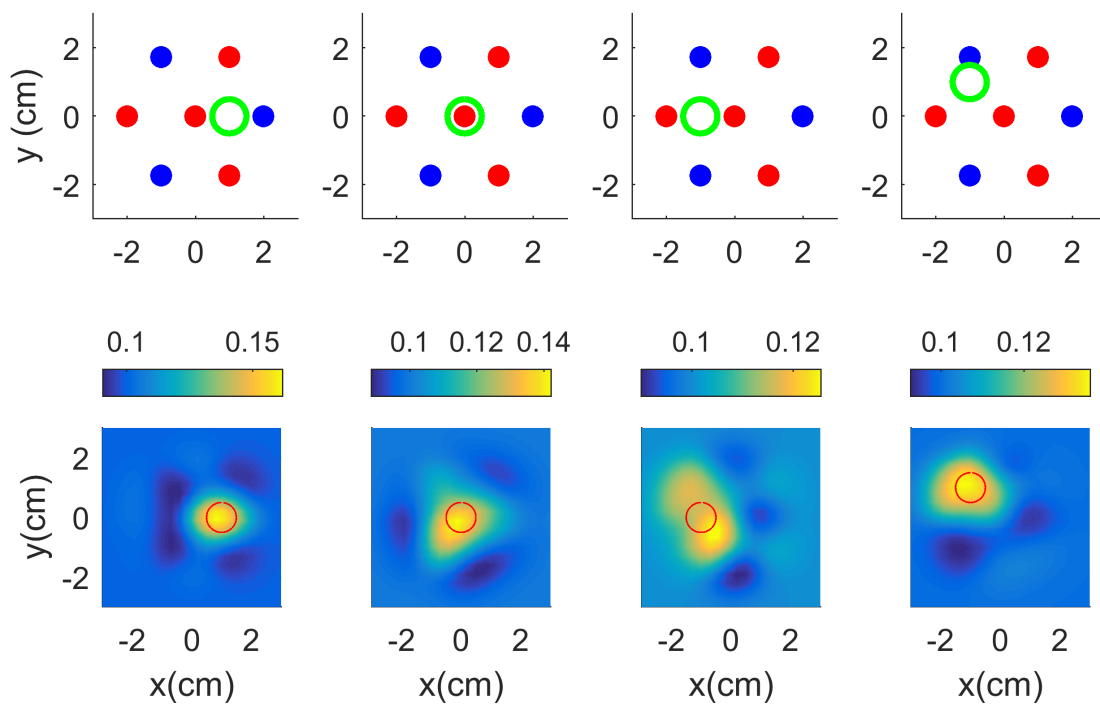


Figure 5.4: Reconstructions for second probe prototype. Colorbars quantify the reconstructed absorption.

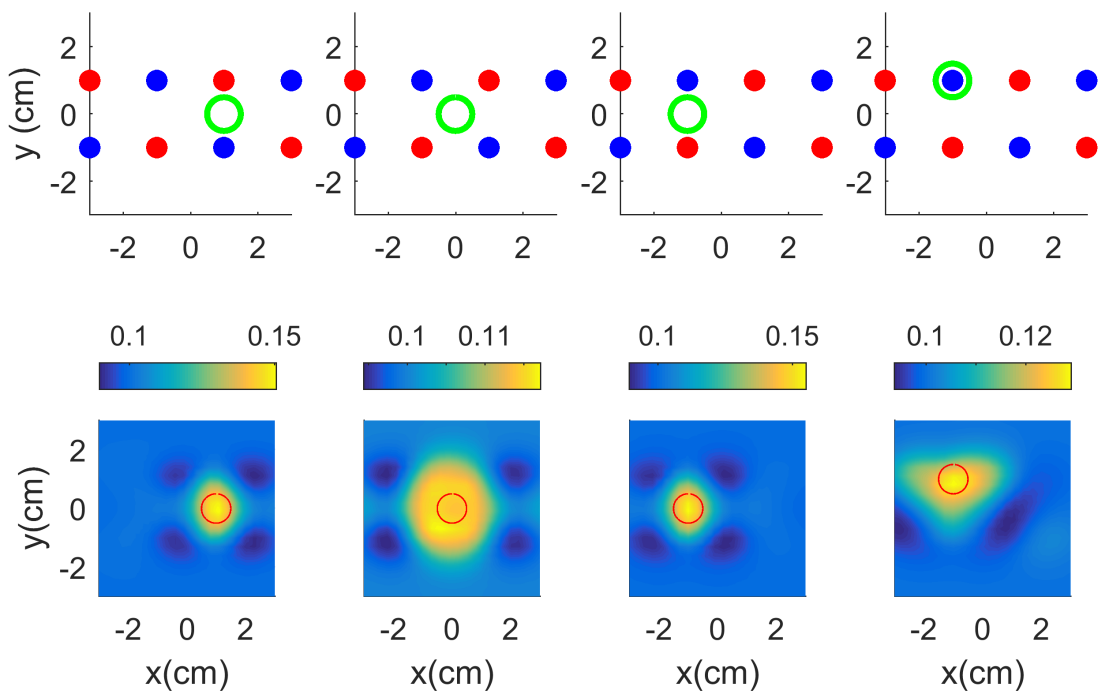


Figure 5.5: Reconstructions for third probe prototype. Colorbars quantify the reconstructed absorption.

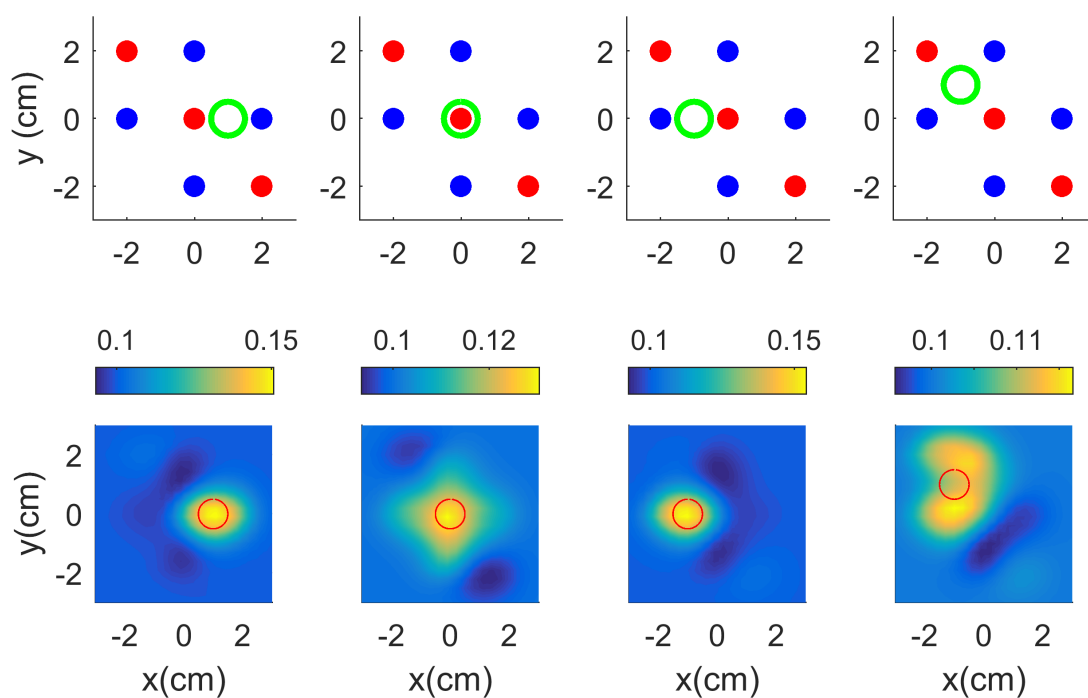


Figure 5.6: Reconstructions for fourth probe prototype. Colorbars quantify the reconstructed absorption.

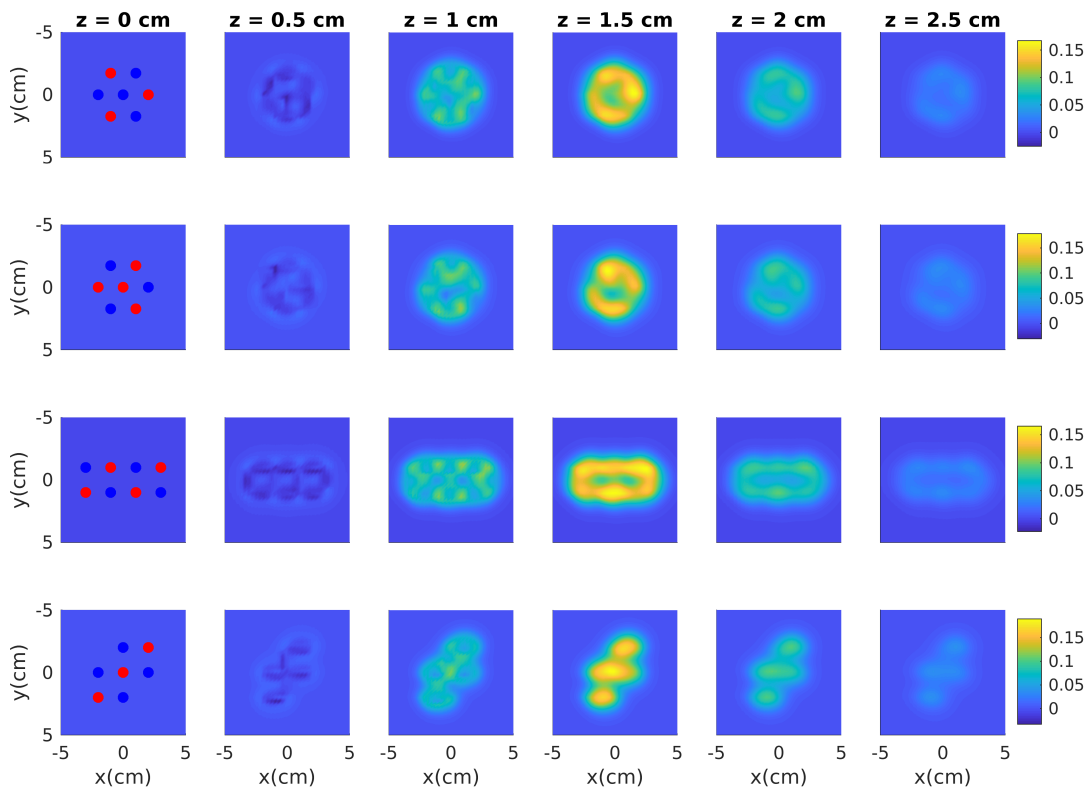


Figure 5.7: Reconstructions for each prototype where a  $\Delta\mu_a = 0.2 \text{ cm}^{-1}$  and 0.5 cm thick layer was located in between 1 to 1.5 cm depth.



Figure 5.8: (Left) One of the probes located at the surface of a solid switchable phantom. (Right) Probe located at the arm of one of the subjects during arm occlusion experiment. Green and black probes are sources and detectors respectively.

## 5.2. Methods

In this section, I will describe the data preprocessing, analysis and tomographic reconstructions that I used to process the data obtained from the experiments.

### 5.2.1. Data preprocessing

Before starting with data preprocessing part, I would like to define some terminology. The in-vivo experiments were divided in several *repetitions* or *blocks*, that is, for arm occlusion experiments each protocol was repeated three times and for motor cortex activation experiments there were done five repetitions, see Sections 5.3 and 5.4 respectively. The experiments were repeated to improve the signal-to-noise ratio and to validate the reproducibility of the physiological phenomenon measured. For each experiment, the instrument performed several *cycles* of 0.96 s duration. Within each cycle, 16 DTOFs were measured which were averaged as explained in the next paragraph.

For this system, data preprocessing is a critical step to perform before doing data analysis or tomography. The first issue to be addressed is the fact that light switching is done automatically by a switcher which it is not synchronized with the detectors, that is, in the preprocessing it has to be detected when the switching occurred. In Figure 5.9 (Left), the distributed time of flight (DTOF) curves measured at a given detector at a given cycle for an arm occlusion experiment are shown. It can be seen that first two curves (blue and red) are background noise and the switching started after second curve was measured. The count rate of the third curve (yellow) is lower than the next curves, therefore it was discarded. To be safe, the fourth curve was also discarded. After, the average of remaining twelve curves is performed, see Figure 5.9 (Right), and background noise was subtracted.

The second part of data preprocessing consist on trimming the curves using the 3%-5% rule, that is, curves are trimmed at the left of the maximum peak where number of detected photons are less than 3% of the maximum and from the right when are less than 5%.

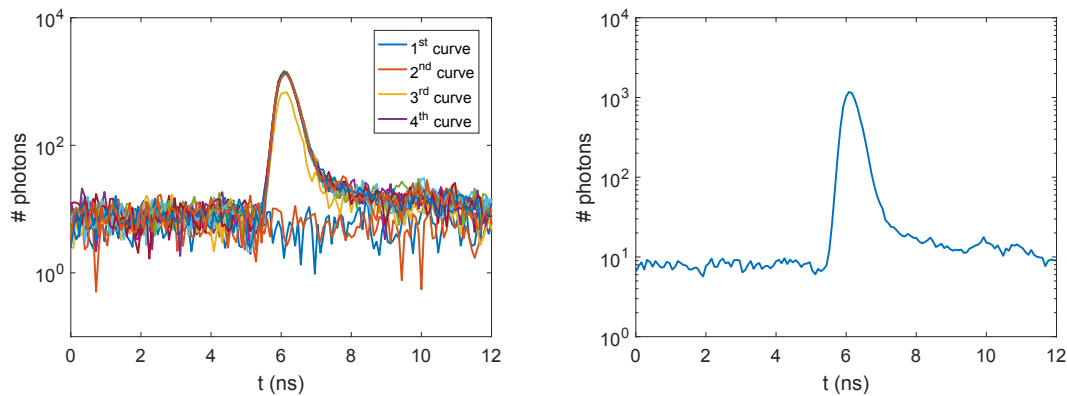


Figure 5.9: (Left) The sixteen measured DTOFs at an arbitrary cycle in an arm occlusion experiment. Switching started after the second DTOF curve; in this case, first four curves are discarded. Rest of the curves are averaged. (Right) Resulting DTOF curve from the average after discarding first four curves.

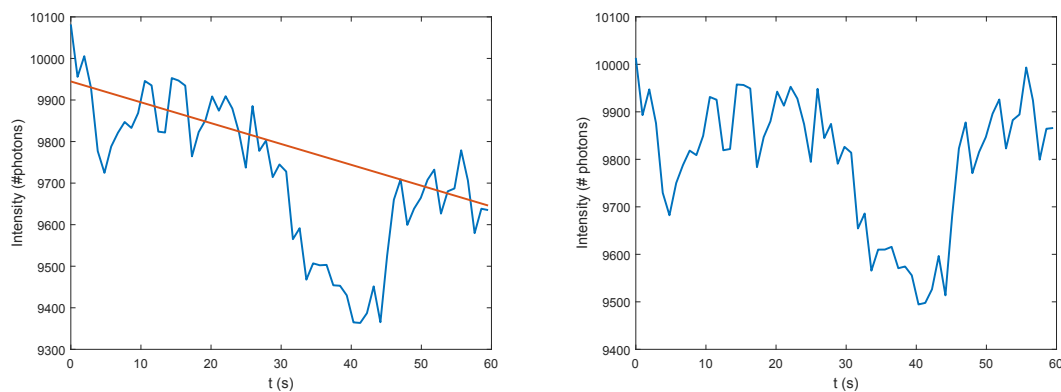


Figure 5.10: (Left) The blue line represents the intensity of DTOF curves at an arbitrary detector for a 60 seconds motor cortex activation experiment where finger tapping was performed between seconds 20 and 40. The red line is the down trend of the intensity represented by fitting a line to the rest and recovery period (in this case 0 – 20s and 50 – 60s). (Right) Intensity of DTOF curves after trending correction using the red line linear approximation.

The third part of data processing takes into account that measurements have also to be detrended in some cases due to thermal drift and other experimental factors. For each repetition the intensity of the DTOF curves is computed and a linear approximation is calculated by fitting the resting and recovery periods, see Figure 5.10 (Left). The resulting detrended intensity is shown at Figure 5.10 (Right).

### 5.2.2. Data analysis

Before performing tomographic reconstructions, I did a preliminary analysis using standard spectroscopy data analysis approaches. With those approaches neither spatial nor depth

information is obtained but they are useful to validate the tomographic reconstruction approaches.

For arm occlusion experiments, I integrated the time-resolved curves to get the continuous-wave (CW) signal. The justification for using CW signals is that forearm muscles are very close to the surface since fat layer is not very thick (just a few millimeters), therefore there is no need of using moments with higher sensitivity to deeper layers.

The absorption for CW signals was recovered using modified Beer-Lambert law,

$$-\frac{1}{\langle L \rangle} \ln \left( \frac{I}{I_0} \right) = \Delta \mu_a, \quad (5.1)$$

where  $\langle L \rangle$  is the differential path length factor. After, the oxygenated and deoxygenated hemoglobin values are obtained using standard procedures explained in Subsection 2.2.1.

Regarding motor cortex activation experiments, I decided to use time-of-flight moment since they are more sensitive to deep layers [127]. The absorption from time-of-flight moment can be obtained using a monolayer version of Equation 15 in [210],

$$\Delta \langle t \rangle = \left( \langle L \rangle \langle t \rangle - \frac{\langle L^2 \rangle}{c} \right) \Delta \mu_a, \quad (5.2)$$

where  $\langle L^2 \rangle = c^2 (\text{Var}(DTOF) - \text{Var}(IRF)) + \langle L \rangle^2$ , being  $c$  the speed of light in the medium,  $IRF$  the instrumental response function and  $\Delta \langle t \rangle = \langle t \rangle - \langle t \rangle_0$  where  $\langle t \rangle_0$  is the time-of-flight during resting period.

### 5.2.3. Tomographic reconstructions

Tomographic reconstruction were performed using algorithms based on Mellin-Laplace moments (see Chapter 2). Reconstructions were also performed with Tukey windows (see Chapter 3) using deconvolution techniques but no improvements could be obtained compared to Mellin-Laplace moments. The reason is that, as was shown in Chapter 3, improvements are only seen for inclusion that are at least 2.5 cm depth. However, as will be shown in the next sections, motor cortex activation happened at a depth in between 1.5 to 2 cm, which is not deep enough to see differences between both methods, see Figure 5.13 for a comparison of Mellin-Laplace and Tukey windows in a phantom with an inclusion at 1.5 cm depth. Since Mellin-Laplace moments are still faster to compute, I decided to use them for the reconstructions that are shown below.

The algorithms were initialized using the optical properties estimated from fitting the DTOF curves corresponding to resting period. For example, for motor cortex activation experiments the curves corresponding to first 15 seconds (3/4 of resting period) were fitted using a semi-infinite halfspace model corrected with the instrumental response function of the system. The algorithm was initialized using the average of absorption and scattering values obtained from the fitting of those DTOFs.

Regarding the discretization of the inverse problem, the linear system was weighted with a diagonal matrix. The components of that matrix were  $1/\sqrt{\text{Var}(M_i)}$  where  $M_i$  represents the value of moment  $i$ -th and  $\text{Var}(\cdot)$  is the variance of moments during resting period. The motivation is to introduce into the system, the uncertainty of the moments due to physiological phenomena not related to the occlusion or motor cortex activation.



For the inverse problem, Laplacian regularization was also used. The justification for using such regularization is that there exist some blind spots in the probe geometry as explained in Subsection 5.1.1. Those blind spots cannot be reconstructed by the algorithm, therefore three-dimensional Laplacian regularization is imposed to force reconstruction of those parts with a smooth solution. To implement Laplacian regularization the techniques described in Chapter 4 were used; note that  $\nabla \cdot \nabla = \nabla^2$  and therefore gradient matrices developed at Chapter 4 can be applied. In this case, total variation regularization does not apply since it will not reconstruct the blind spots but promote piecewise constant solutions. In Figure 5.11 a comparison between using or not Laplacian regularization is shown. One of the biggest difference is that quantification of  $\Delta O_2Hb$  is much smaller when Laplacian regularization is used. This effect occurs because Laplacian operator induces a diffusion of the solution and therefore the quantification of  $\Delta O_2Hb$  is distributed over a larger volume. However, for not Laplacian regularized case, there are more heterogeneities which tend to create larger quantification differences. For example, at 1 cm depth the quantification is as big as  $\Delta O_2Hb = 10$  but for  $z = 1.5$  cm negative values are reached which try to compensate large positive values at shallow layers.

In [38], it was reported the use of Laplacian regularization to retrieve the depth of brain activity using diffuse optical tomography. Nevertheless, they propose to use the regularization term  $\|L \delta \mu_a\|_2^2$  where  $L$  is the Laplacian operator. This term imposes Laplacian regularization in the absorption update,  $\delta \mu_a$ , but not in the solution. For this reason, in this work, I decided to use the regularizing term  $\|L(\mu_a + \delta \mu_a)\|_2^2$  that imposes a Laplacian regularized solution.

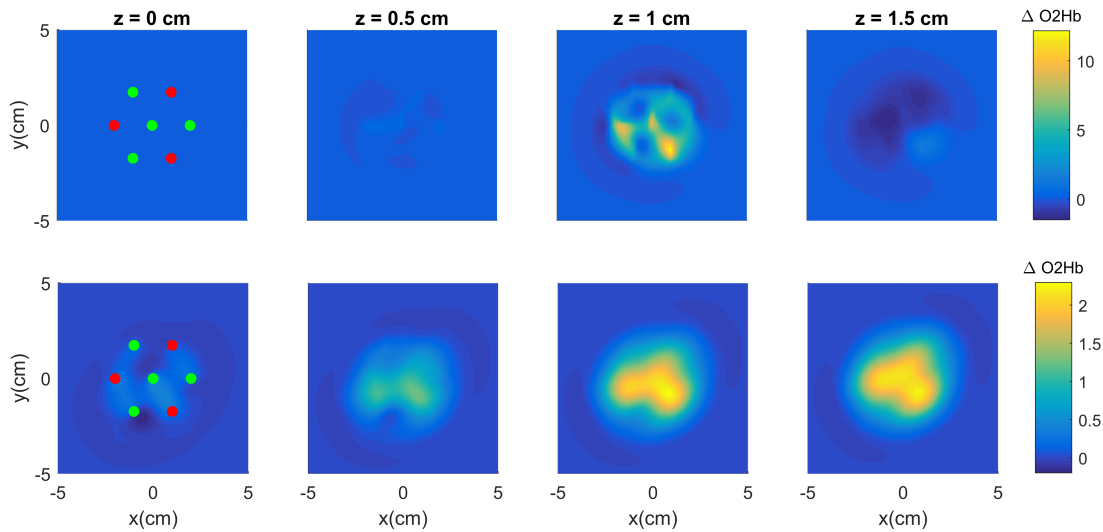


Figure 5.11: Qualitative example of oxygenated hemoglobin reconstruction for a motor cortex activation experiment with and without Laplacian regularization. In the first row, Laplacian regularization was not used and probe blind spots can be clearly seen. In the second row, same case but using Laplacian regularization.

Finally, another critical aspect is to deal with the high sensitivity of shallow layers. A small absorption in surface layers has a large influence in the simulations, therefore the reconstruction algorithm tends to project the absorption reconstructions to the surface. To avoid that effect, the standard practice in the literature is to weight deep nodes to normalize the sen-

sitivity trough all the domain: for example, using a layer-based sigmoid adjustment [254], a radial distance penalty [174, 47] or a depth compensation term [160, 159]. In this work, I normalised the sensitivity of nodes by weighting them with a value proportional to the distance to the surface. The proportionality hyperparameter was selected by performing reconstruction for measurements performed in a switchable phantom [170] with a  $\Delta\mu_a = 0.16 \text{ cm}^{-1}$  inclusion at a 1.5 cm depth, the reconstruction is shown at Figure 5.13. The proportionality hyperparameter was also validated in simulations giving accurate localization results for inclusions up to 2 cm deep, see Figure 5.12. For inclusions at  $z = 2.5 \text{ cm}$  the reconstruction is projected to  $z = 2 \text{ cm}$  (see last row from Figure 5.12).

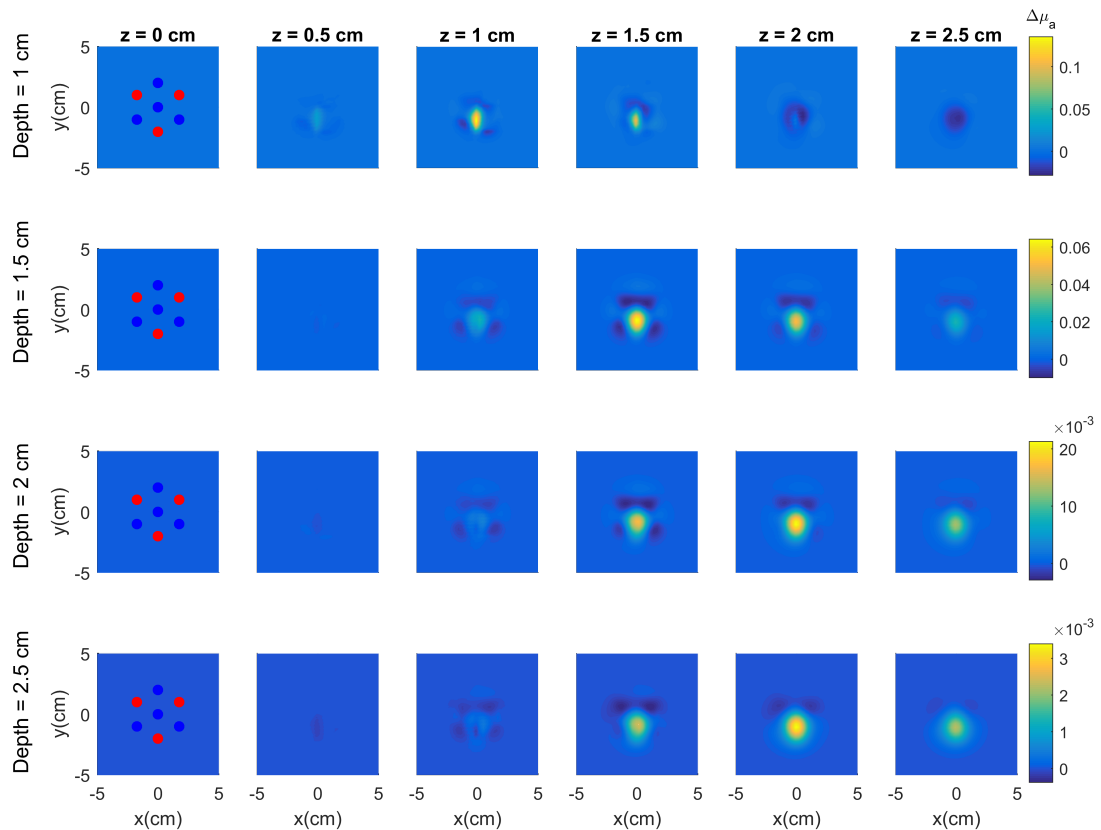


Figure 5.12: Reconstructions from simulations of a cylindrical inclusion of 5 mm diameter, 5 mm length located at  $x = (0, -1, 1 \text{ to } 2.5) \text{ cm}$  with  $\Delta\mu_a = 0.2 \text{ cm}^{-1}$ . Between 1 cm to 2 cm depth the reconstruction locates precisely the inclusion.

Nevertheless, even normalizing the sensitivity still shallow layers will affect considerably the reconstruction. For human brain measurements, this effect is critical since hemodynamics are constantly changing in the scalp tissue and they contaminate the brain activation signals. To avoid these issues, the approach that I took in this work, was to reconstruct the domain from bottom to the surface. This method works as follows: in each iteration, reconstruction is performed only from a given depth value to the bottom. For example, for a  $z_{\max} = 5 \text{ cm}$  cubic domain with a starting value of 0.4 cm depth and an increase of 0.2 cm, the reconstruction iterations will be: (1) at first iteration the reconstruction is performed for nodes between 0.4 cm and 5 cm depth, (2) at second iteration the reconstruction is performed for nodes between 0.2 cm and 5 cm, (3) at third and next iterations the reconstruction

is performed in all the domain until convergence is reached. This approach forces the algorithm, during the first iterations, to perform the reconstruction at deeper layers without using superficial nodes. Up to my knowledge, this approach was not reported before in the literature. In the following reconstruction, an initial starting value of 0.8 cm was used and steps were of 0.2 cm thickness. Therefore, at fifth iteration the algorithm starts to reconstruct the whole domain.

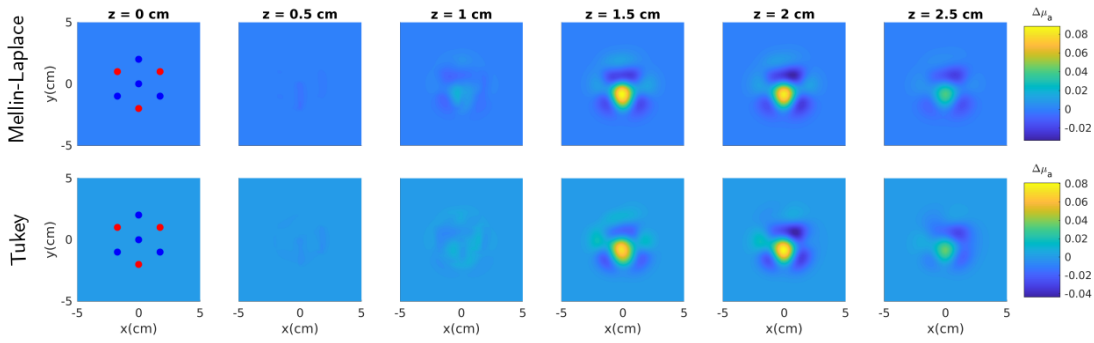


Figure 5.13: Reconstructions of a cylindrical inclusion of 5 mm diameter, 5 mm length located at  $x = (0, -1, 1.5)$  cm depth inside a solid phantom. The equivalent absorption for 670 nm was  $\Delta\mu_a = 0.16 \text{ cm}^{-1}$ ; for further details refer to [170]. (First row) Reconstruction performed Mellin–Laplace moments. (Second row) Reconstruction performed using Tukey windows. No significant differences are seen.

### 5.3. Arm occlusion experiments

Arm occlusion experiments are easy to perform and the physiological process behind them is well known. Therefore, these experiments are ideal to use as a validation test. Two different types of arm occlusion experiments were performed: venous and arterial occlusion. The difference stands in the pressure that was input to the cuff and the physiological consequences that it has. The occlusion tests were performed to three different subjects (two men and one female).

The vasculature of human arm is composed of arteries and veins, see Figure 5.14. Meanwhile, the veins are closer to the surface of the arm the arteries are deeper. Arteries are in charge of transporting oxygenated blood to the arm. After, the muscles take the oxygenated blood from the arteries and consume the oxygen molecules transforming the blood to a deoxygenated state. Finally, the veins carry the deoxygenated blood to the heart.

In the next subsections, I will explain the physiological process that takes place in venous and arterial occlusion and I will show the data analysis and tomography reconstructions I performed.

#### 5.3.1. Venous occlusions

The physiological process during a venous occlusion goes as follows: the pressure of the cuff is increased up to levels slightly below the systolic pressure of the subject. For the three

subjects the systolic pressure was between 120 – 130 mmHg and the arm cuff pressure was increased up to 100 mmHg. When cuff pressure is not increased above systolic pressure only veins are blocked since they are closer to the surface. However, arteries are not blocked because the cuff pressure is not large enough. Therefore, the arm output blood flow is blocked meanwhile the arteries are still open. Hemoglobin quantity at the occluded arm will increase since blood can enter the arm but not escape. In overall, both oxy/deoxygenated hemoglobin will increase in the arm during the occlusion but oxygenated hemoglobin quantity will be higher than deoxygenated since the rate of consumption of oxygenated blood by the arm is lower than the oxygenated input blood flow from arteries.

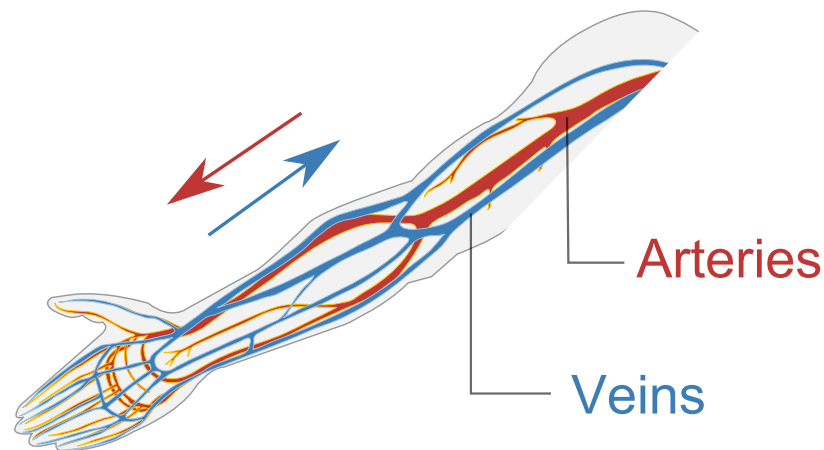
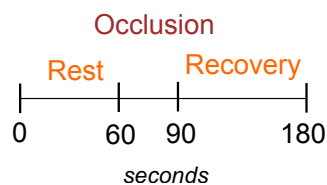


Figure 5.14: Arteries and veins in the human arm. Arteries/veins carry oxygenated/deoxygenated blood toward/away from the arm. (Modified figure taken from Wikimedia Commons).

The venous occlusion protocol consisted in 60 s of resting, 30 s of occlusion at 100 mmHg and finally 90 s of recovery. The protocol was repeated three times, see Figure 5.15 (Left).

#### Venous occlusion protocol



#### Arterial occlusion protocol

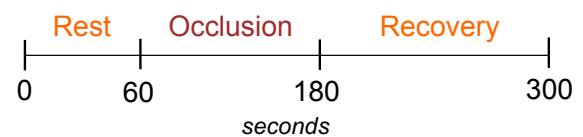


Figure 5.15: (Left) Venous occlusion protocol. (Right) Arterial occlusion protocol. Both protocols were repeated three times for all the subjects.

In Figure 5.16 it is shown where the probes were located for all subjects. The orientation of the probes was the same for all subjects and the cuff was always wore at the left arm.

For all arm occlusion experiments, CW data analysis was used to obtain the time series results. In Figure 5.17, the time series of oxygenated and deoxygenated blood for Subject 1

occluded arm during venous occlusion experiment is shown. As expected, both oxy and deoxygenated blood are increasing during the occlusion (60 s – 90 s); increase of oxygenated blood is higher because the rate of consumption of muscles is slower than the input flow of oxygenated blood. At some source–detector couples (e.g. 2–5, 4–5, 4–7,4–8) the slope of the increase is steeper, such differences can be explained because the inhomogeneous presence of veins and arteries in the arm. After occlusion, oxy/deoxygenated blood levels return to preocclusion values quickly. The values at reference arm (see Figure 5.18) are stable along the whole experiment. Similar conclusions are reached regarding Subject 3, see Figures 5.21 and 5.22.

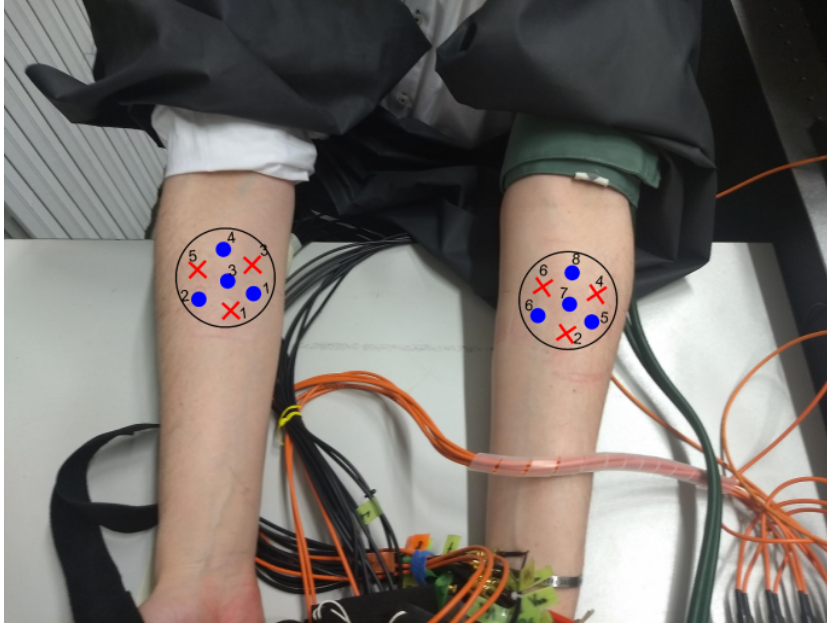


Figure 5.16: Location of the two probes during arm occlusion experiments. The probes were equally orientated for all the subjects. Notice that the green cuff is located at the left arm. Right arm was consider as reference.

For Subject 2, the Figure 5.19 shows an anormal behaviour of hemoglobin values after realising the cuff pressure. This behaviour occurred in the three repetitions. It is also interesting to notice that for reference arm, see Figure 5.20, abnormal variation of oxygenated blood happened after occlusion times. These results suggest that the cuff release induced a physiological response in the subject body which altered the hemoglobin values in the occluded arm for one minute – after sixty seconds the values return to pre-occlusion levels.

From the results, it is also clear that  $\Delta O_2Hb$  measurements have larger uncertainty, that is, the standard deviation is greater. This phenomenon is explained by the specific absorption matrix associated to 670 nm and 820 nm wavelengths,

$$A = \begin{bmatrix} 6.43 & 0.67 \\ 1.59 & 2.1 \end{bmatrix} \times 10^{-3} \quad [\text{cm}^{-1} \mu M] \quad (5.3)$$

where first and second row correspond to specific absorption values for 670 nm and 820 nm wavelengths for deoxy and oxygenated hemoglobin. Let assume that the covariance of our

measurements is independent and equal, that is, the covariance matrix  $M$  is,

$$M = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (5.4)$$

When transforming the measurements to oxy and deoxygenated hemoglobin values by solving the linear system (see Subsection 2.2.1) the new covariance matrix  $M'$  is,

$$M' = ((A^T A)^{-1} A^T) M ((A^T A)^{-1} A^T)^T = \begin{bmatrix} 0.31 & 0.49 \\ 0.49 & 2.84 \end{bmatrix}. \quad (5.5)$$

Therefore, from the previous covariance matrix it can be deduced that the uncertainty of  $O_2Hb$  will always be greater than the uncertainty of  $HHb$ .

The tomographic reconstructions videos for venous occlusions can be found at [74]. The videos were obtained by reconstructing each set of measurements and averaging over the three repetitions, that is, at a given time the reconstructions were performed for the three repetitions and the results were averaged. The reconstructions clearly validate the data analysis. Moreover, it can also be seen that most of changes occur at surface tissue, that is, between layers from 0 to 1 cm.

For Subject 1, reconstructions show a huge increase of hemoglobin concentration at positions where source 2 and 4 are located. The increase of deoxygenated blood is also seen at the video but the color intensity is around  $8 \mu M$  lower. These results are in accordance with Figure 5.17.

Subject 2 presents a normal increase of  $O_2Hb$  concentration although it is also seen an abnormal decrease of  $HHb$ . This abnormal behaviour could be a consequence of the hemoglobin concentration instabilities that were observed for this subject.

For Subject 3, reconstruction are more localized at positions of sources 2 and 6, as expected from Figure 5.21. In this case, maximum values were located at 1 cm depth, that is, deeper than for Subject 1. A possible explanation is that this subject had a thicker surface skin and fat layer in the arm.

### 5.3.2. Arterial occlusions

For arterial occlusions the cuff pressure is at least twice higher than the systolic pressure of the subjects. In this case, veins and arteries are blocked, therefore the blood inflow and outflow is totally blocked at the arm. Ideally, hemoglobin quantity will be constant during the occlusion. In overall, oxygenated hemoglobin quantity will decrease since it is consumed by the arm muscles and it will be transformed to deoxygenated hemoglobin.

The arterial occlusion protocol consisted in 60 s of resting, 120 s of occlusion at 250 mmHg and finally 120 s of recovery, see Figure 5.15 (Right).

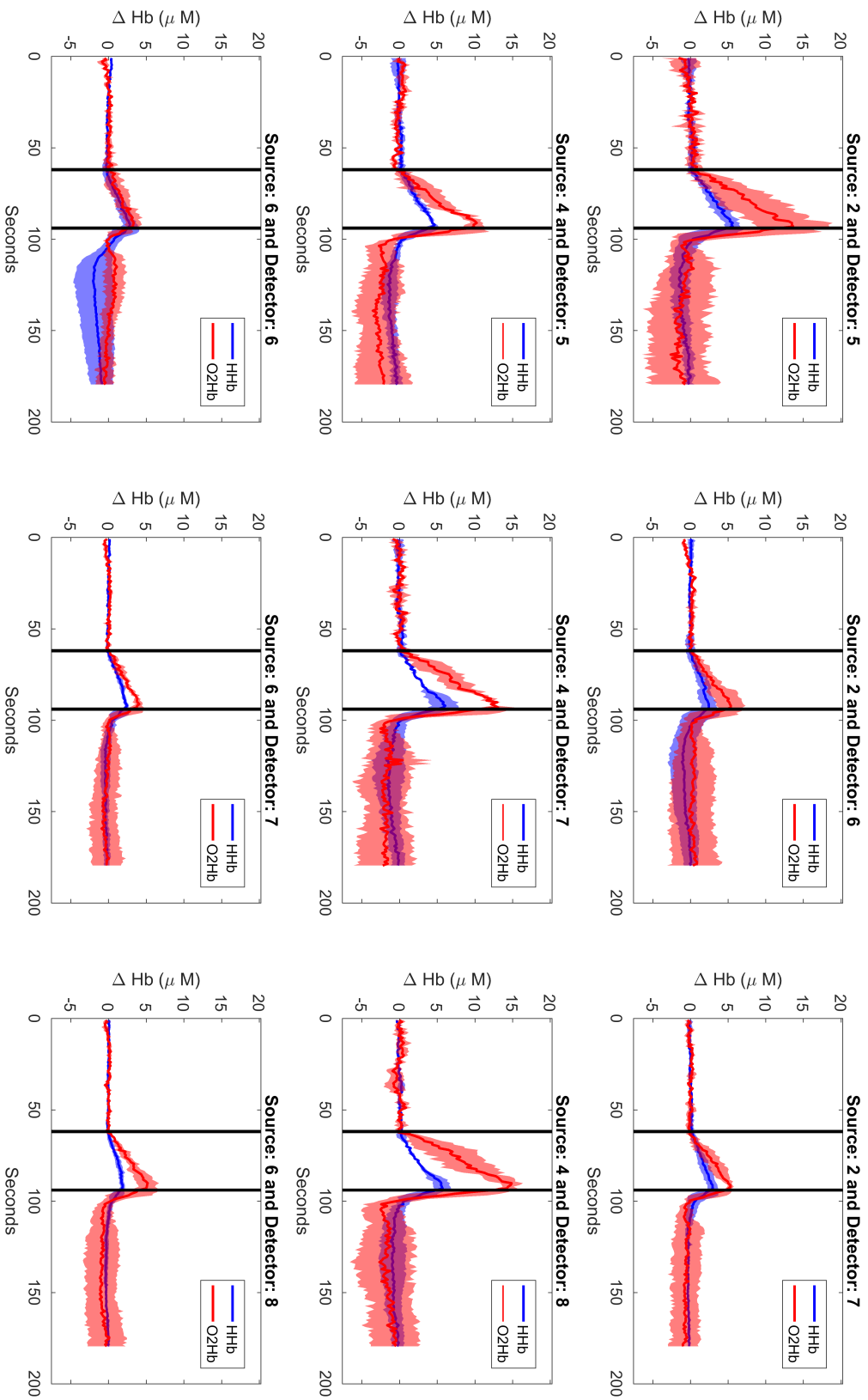


Figure 5.17: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for occluded arm during venous occlusion experiment performed at Subject 1. The shadows indicate the standard deviation over three repetitions. Occlusion happened between black vertical lines.

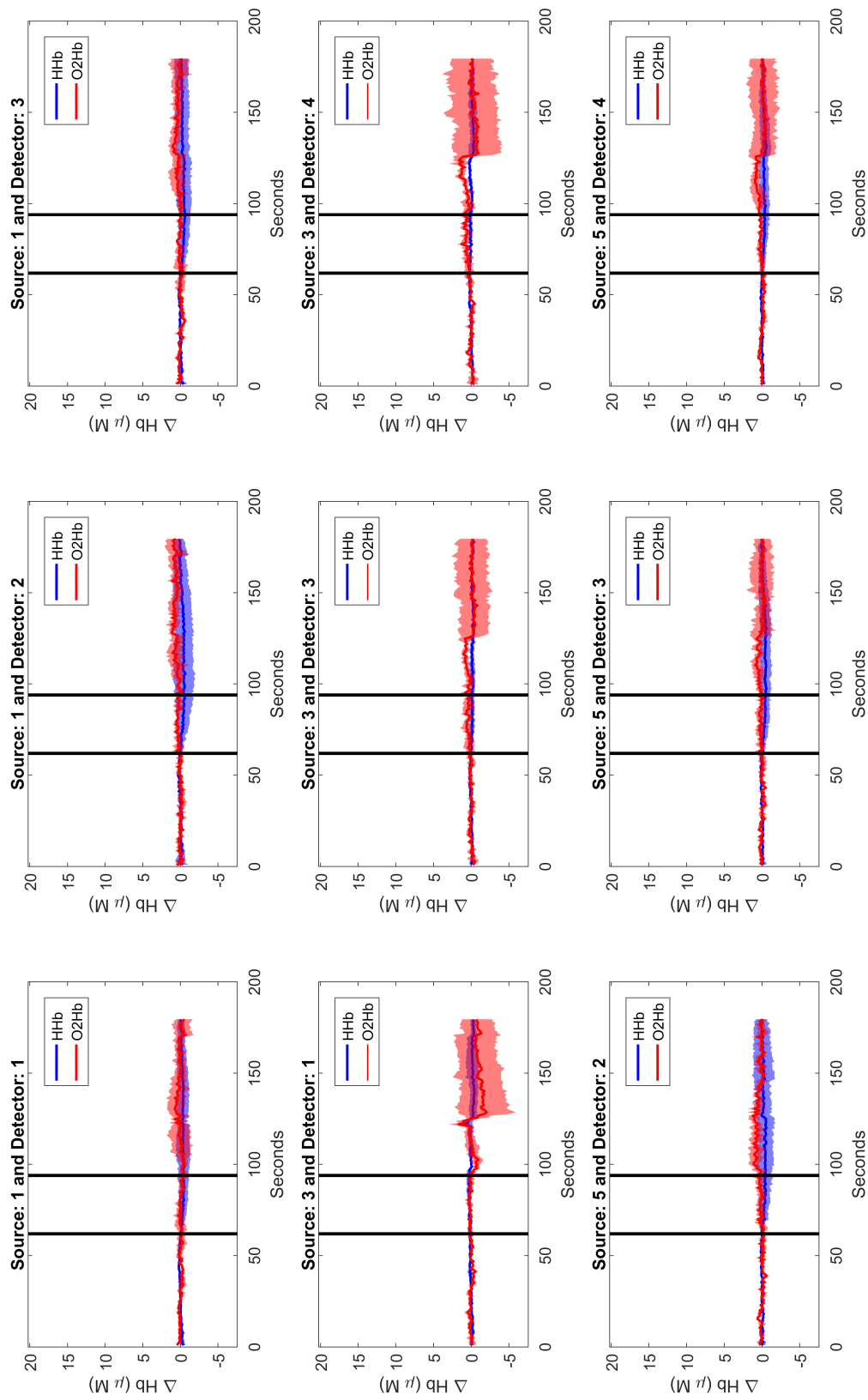


Figure 5.18: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for reference arm during venous occlusion experiment performed at Subject 1.



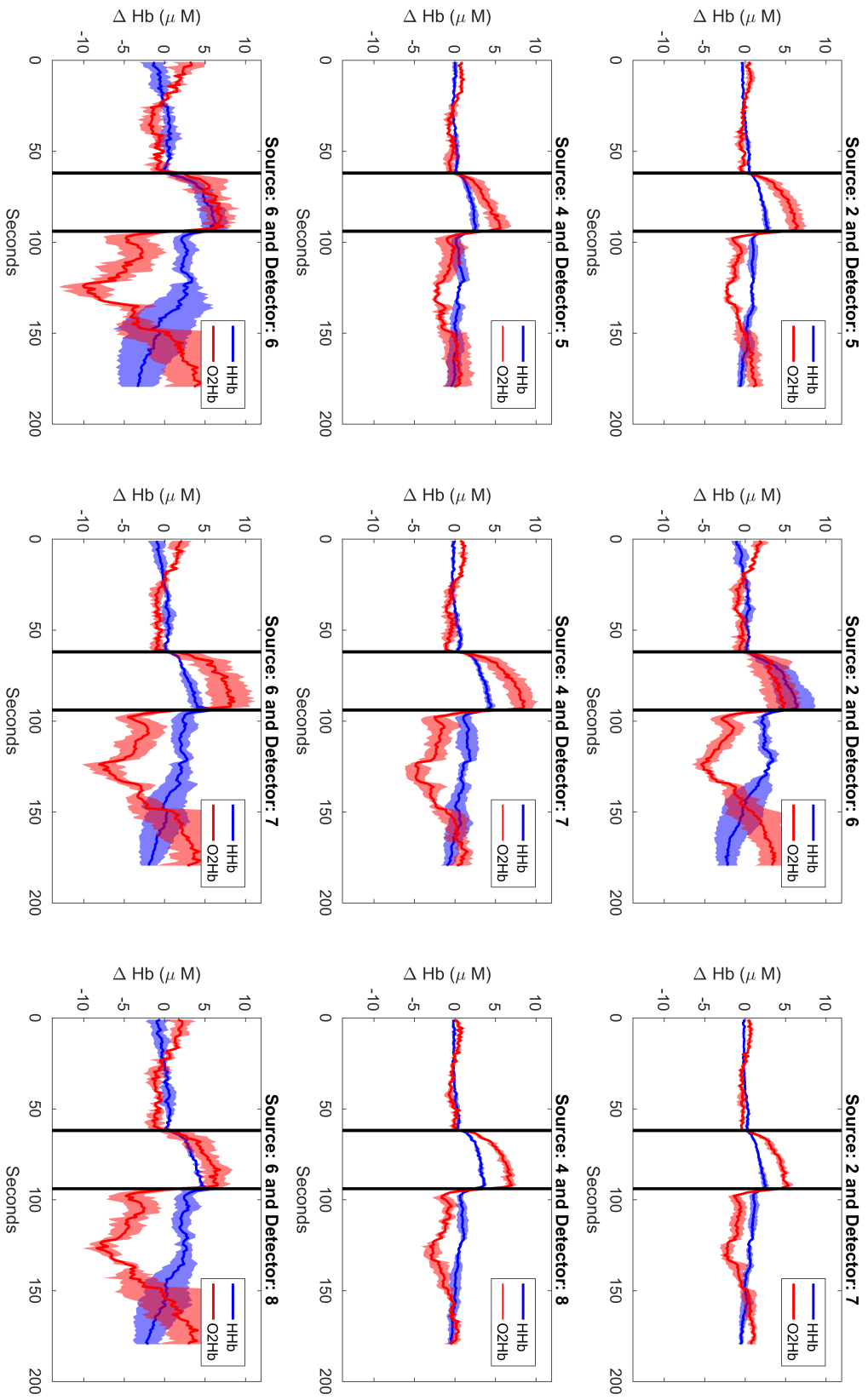


Figure 5.19: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for occluded arm during venous occlusion experiment performed at Subject 2.

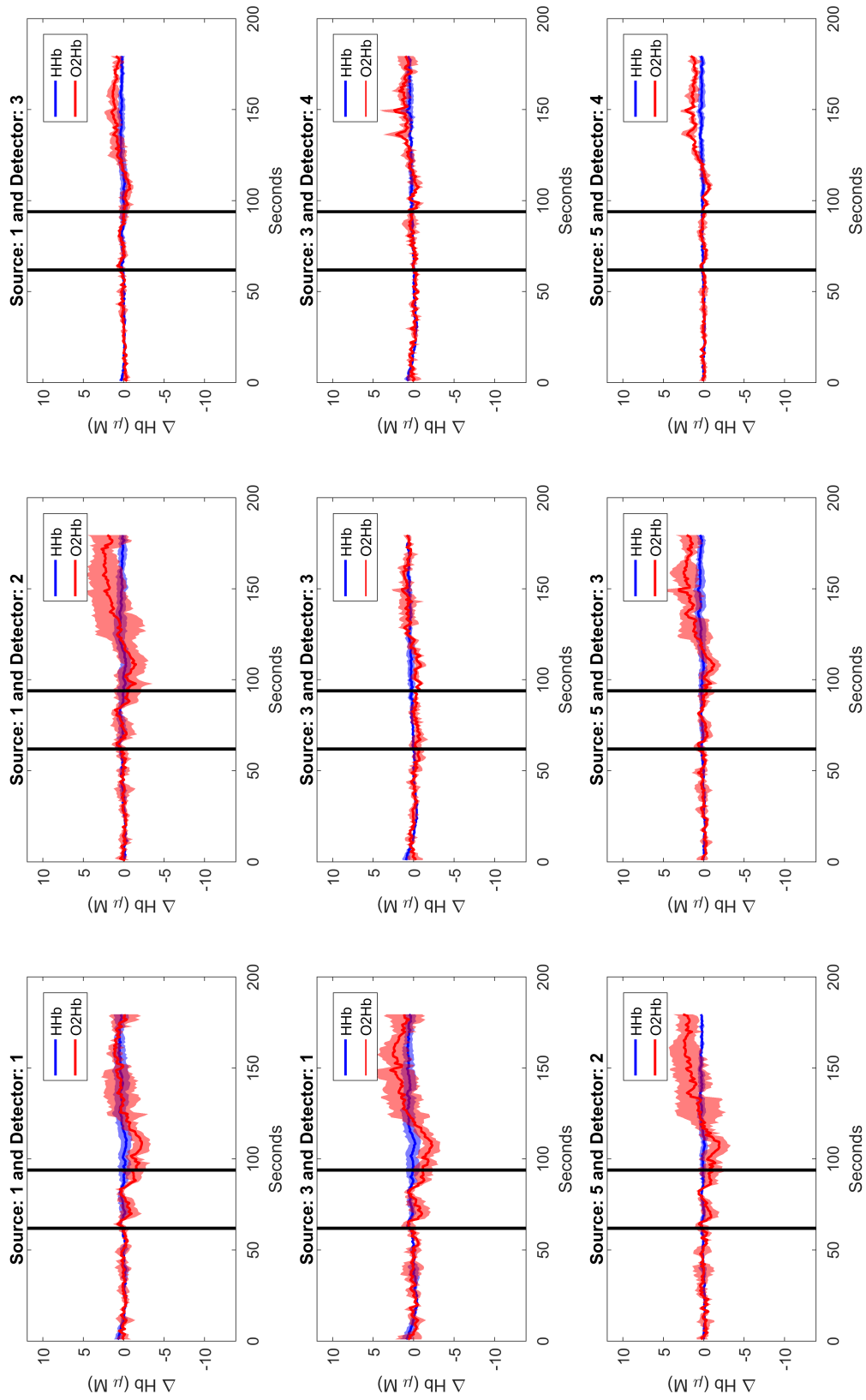


Figure 5.20: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for reference arm during venous occlusion experiment performed at Subject 2.

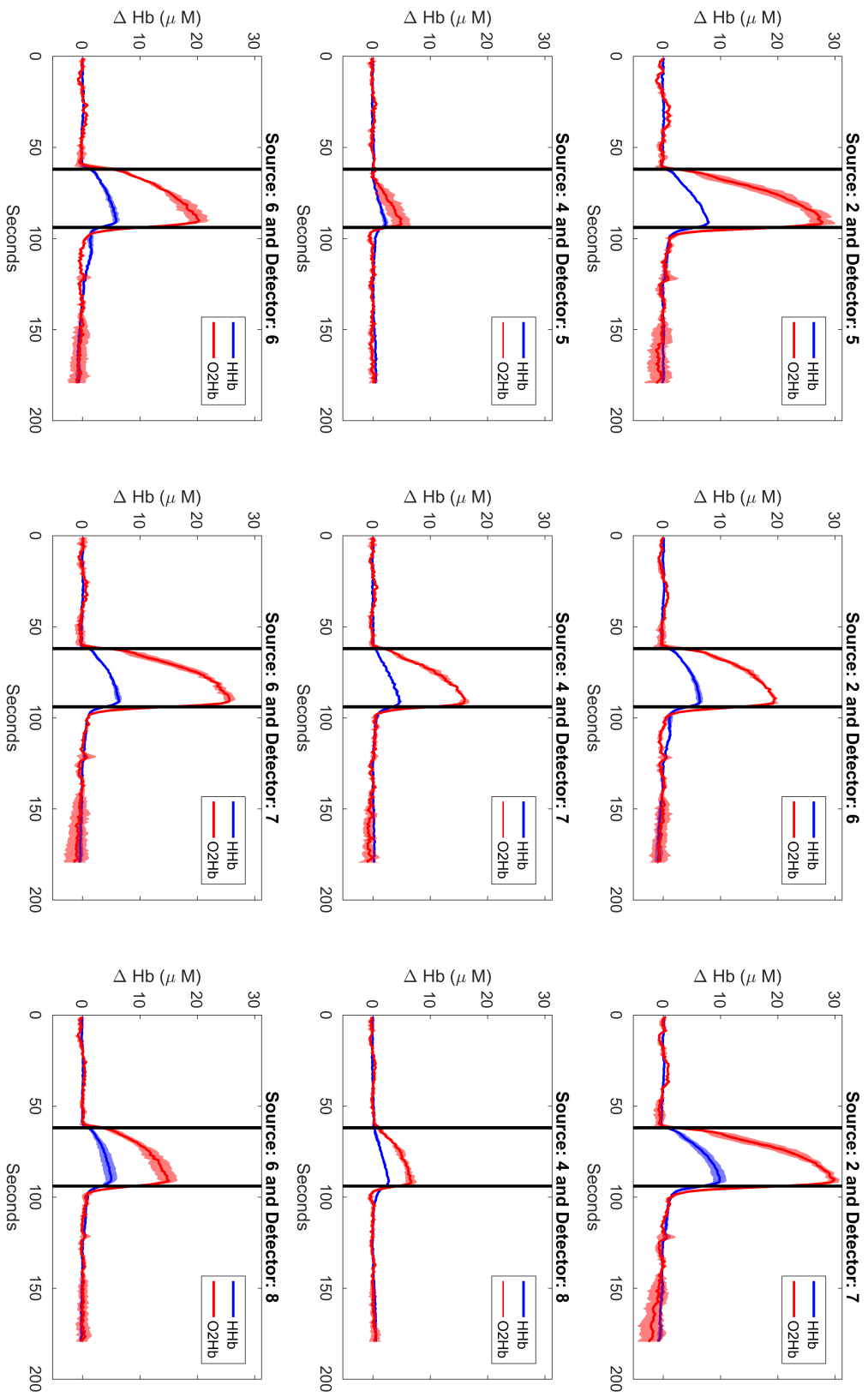


Figure 5.21: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for occluded arm during venous occlusion experiment performed at Subject 3.

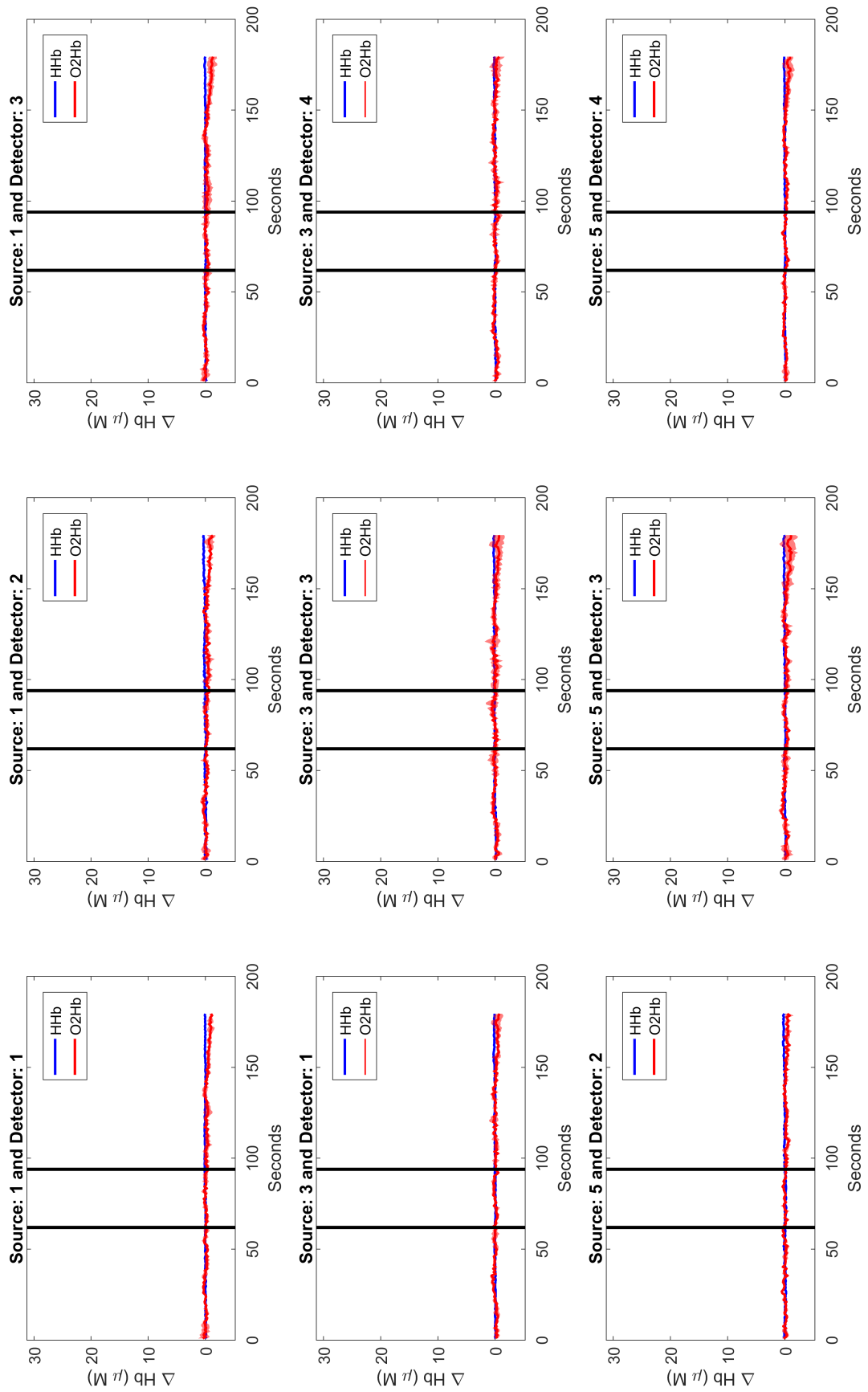


Figure 5.22: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for reference arm during venous occlusion experiment performed at Subject 3.

In Figure 5.23, the time series of oxygenated and deoxygenated hemoglobin for Subject 1 occluded arm during arterial occlusion experiment is shown. The results are in accordance with the physiological behaviour explained before; during arterial occlusion the deoxygenated blood increases and oxygenated blood decreases at equal rates. Moreover, the total hemoglobin is maintained constant. Few seconds after releasing cuff pressure, the concentration of oxygenated blood increases dramatically since arteries are suddenly opened. Similar phenomenon happens with deoxygenated blood but in the opposite way, that is, the concentration of deoxygenated blood dramatically decreases. After one minute, the concentrations of oxygenated blood return to baseline values; deoxygenated blood also tends to recover baseline values although still needs a few more seconds. It must be noted that for some channels (e.g. 2–5, 4–5, 4–7 or 4–8) at the beginning of the occlusion the oxygenated blood increases briefly. This happens because the increase of cuff pressure to 250 mmHg value is not immediate and therefore at the beginning a venous occlusion takes place (note that in this case the total hemoglobin also increases). This effect only happens during the first 10 seconds after occlusion starts. For values at reference arm, see Figure 5.24, a slight increase/decrease of oxygenated/deoxygenated blood close to the end of the occlusion is seen but is not significant.

For Subject 2, see Figure 5.25, the increase/decrease of deoxy/oxygenated hemoglobin is also observed. However, the variability of the hemoglobin quantification is greater depending on the channel (e.g. compare channels 2–5 and 6–6). It is also seen the same phenomenon after releasing cuff occlusion as for venous occlusion.

The Figure 5.27 shows the times series for Subject 3 during arterial occlusion experiments. For channels 4–5, 4–7 and 4–8 the results are very similar to Subject 1. For the rest of the channels, it can be seen a huge increase of oxygenated blood during the first five seconds after occlusion started. The explanation is the same as given for previous subjects – when cuff pressure is increased the 250 mmHg value is not reached immediately and during the first seconds venous occlusion is also induced. Nevertheless, after the first five seconds the decrease of oxygenated blood and increase of deoxygenated blood is as expected. The values for reference arm are stable during the experiments, see Figure 5.28.

The tomographic reconstructions videos for arterial occlusions are found at [74]. These reconstructions also validate the data analysis. For Subject 1, the increase of deoxygenated hemoglobin concentration during the occlusion and the decrease of oxygenated hemoglobin concentration is clearly seen. As for venous occlusions, higher concentration variabilities are seen at positions close to source 2 and 4. Moreover, after releasing the cuff pressure, the concentration values swap rapidly and during the recovery period they approach  $\Delta\text{Hb} = 0$ . Although, for deoxygenated blood preocclusion values are not totally recovered as explained in the time-series analysis.

For Subject 2, results are less clear because the increase of oxygenated hemoglobin concentration during the first seconds of occlusion. However, the expected decrease of oxygenated hemoglobin concentration during occlusion is seen. The deoxygenated hemoglobin concentration increase is barely seen.

For Subject 3, an interesting observation is seen for oxygenated blood: at positions of sources 2 and 6, it increases during firsts seconds and for positions close to source 4 it decreases. This phenomenon (explained in time-series data analysis section) is in agreement with Figure 5.27. The oxygenated hemoglobin decreases after the first seconds of occlusion. Again, for Subject 3 the reconstruction is deeper.

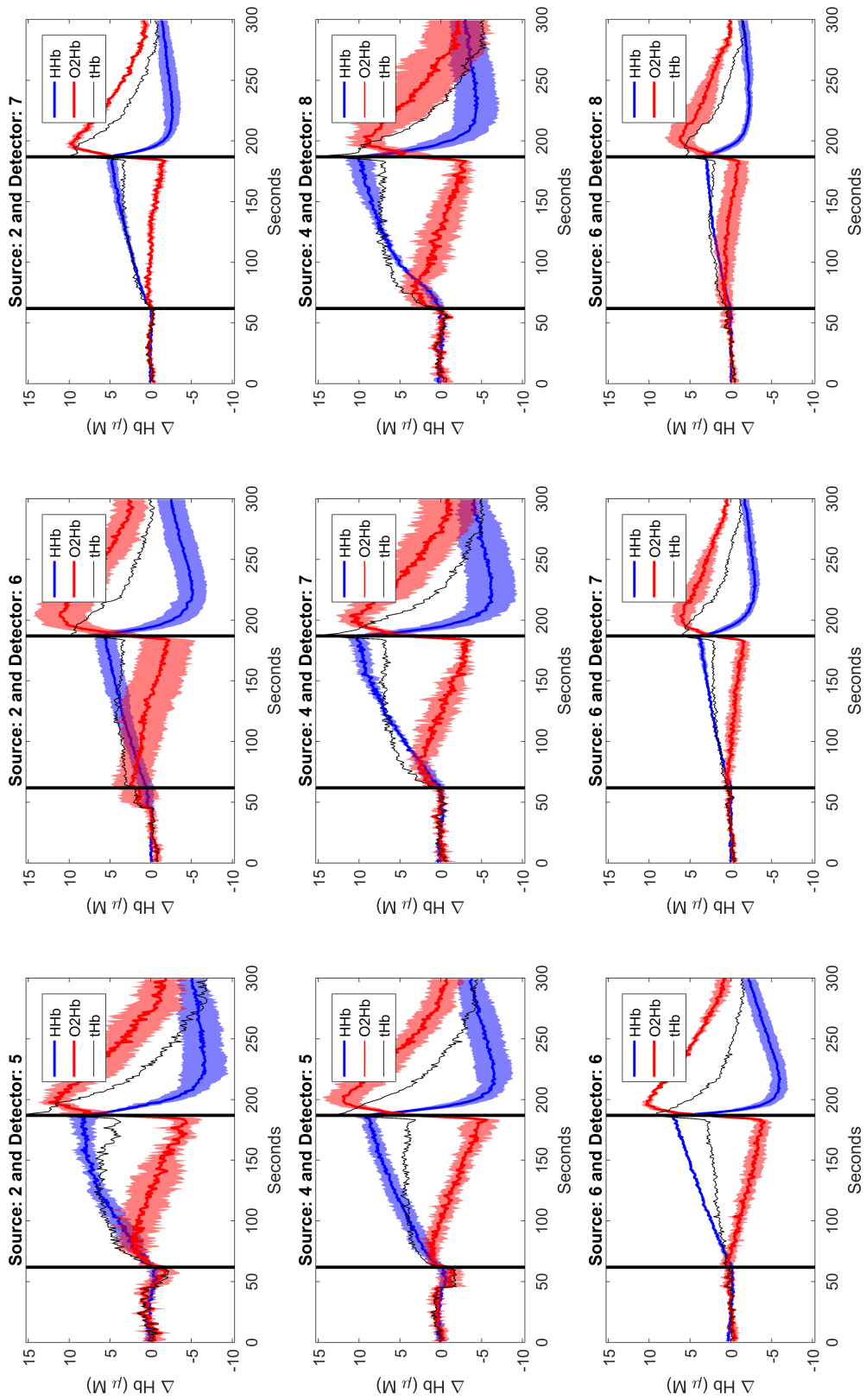


Figure 5.23: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for occluded arm during arterial occlusion experiment performed at Subject 1. Black time series indicates the total hemoglobin value. The shadows indicate the standard deviation over three repetitions. Occlusion happened between black vertical lines.

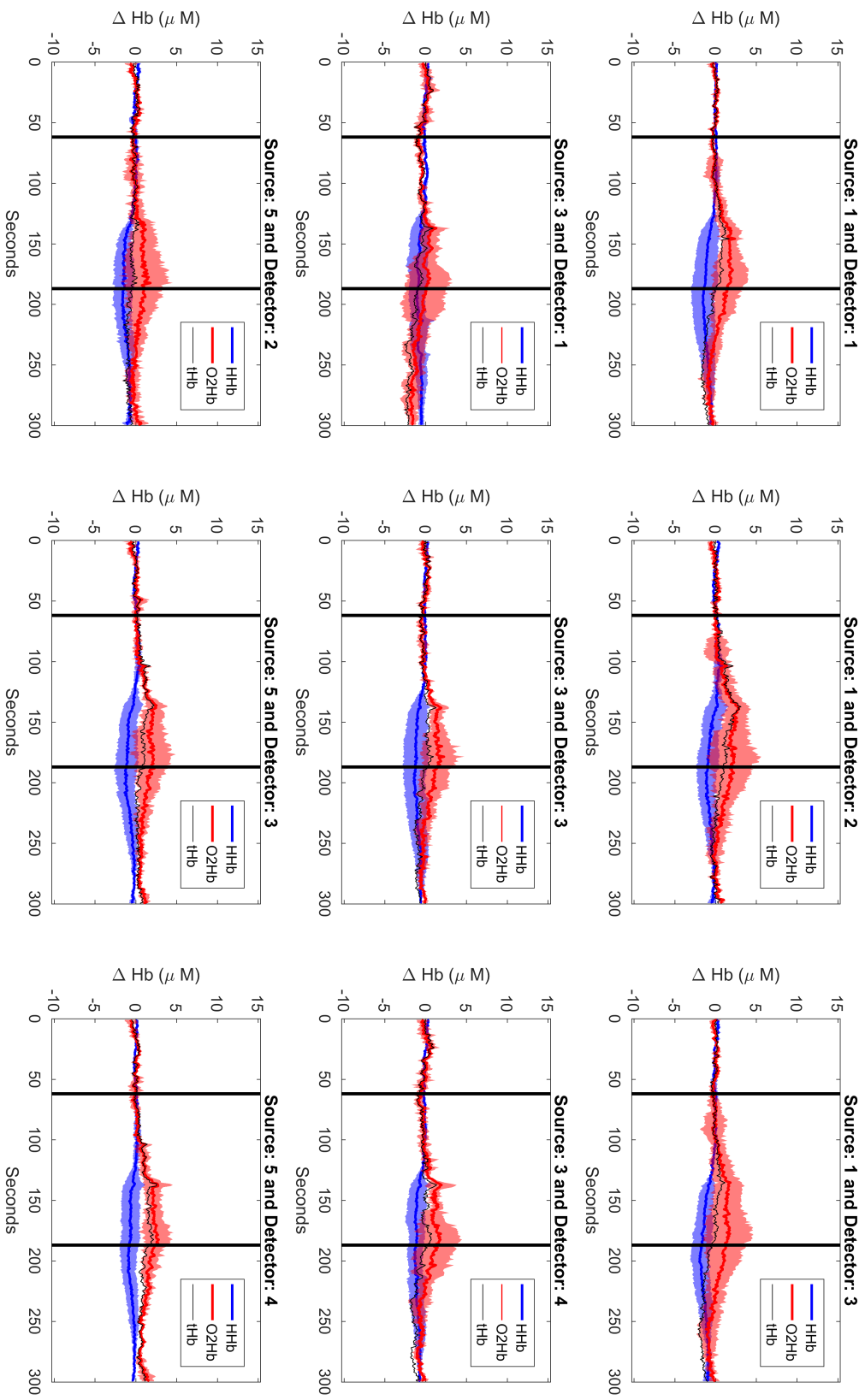


Figure 5.24: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for reference arm during arterial occlusion experiment performed at Subject 1.

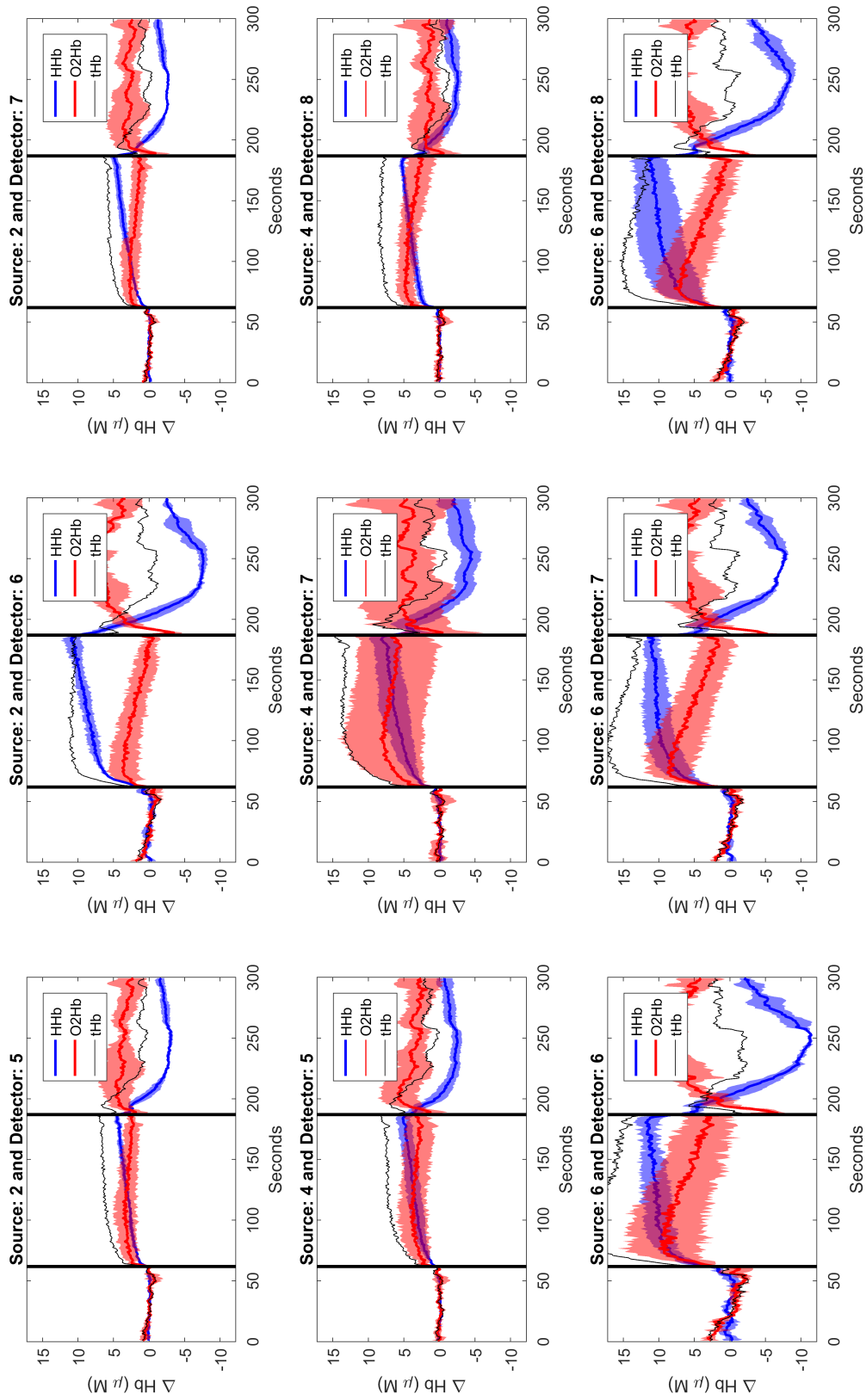


Figure 5.25: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for occluded arm during arterial occlusion experiment performed at Subject 2.



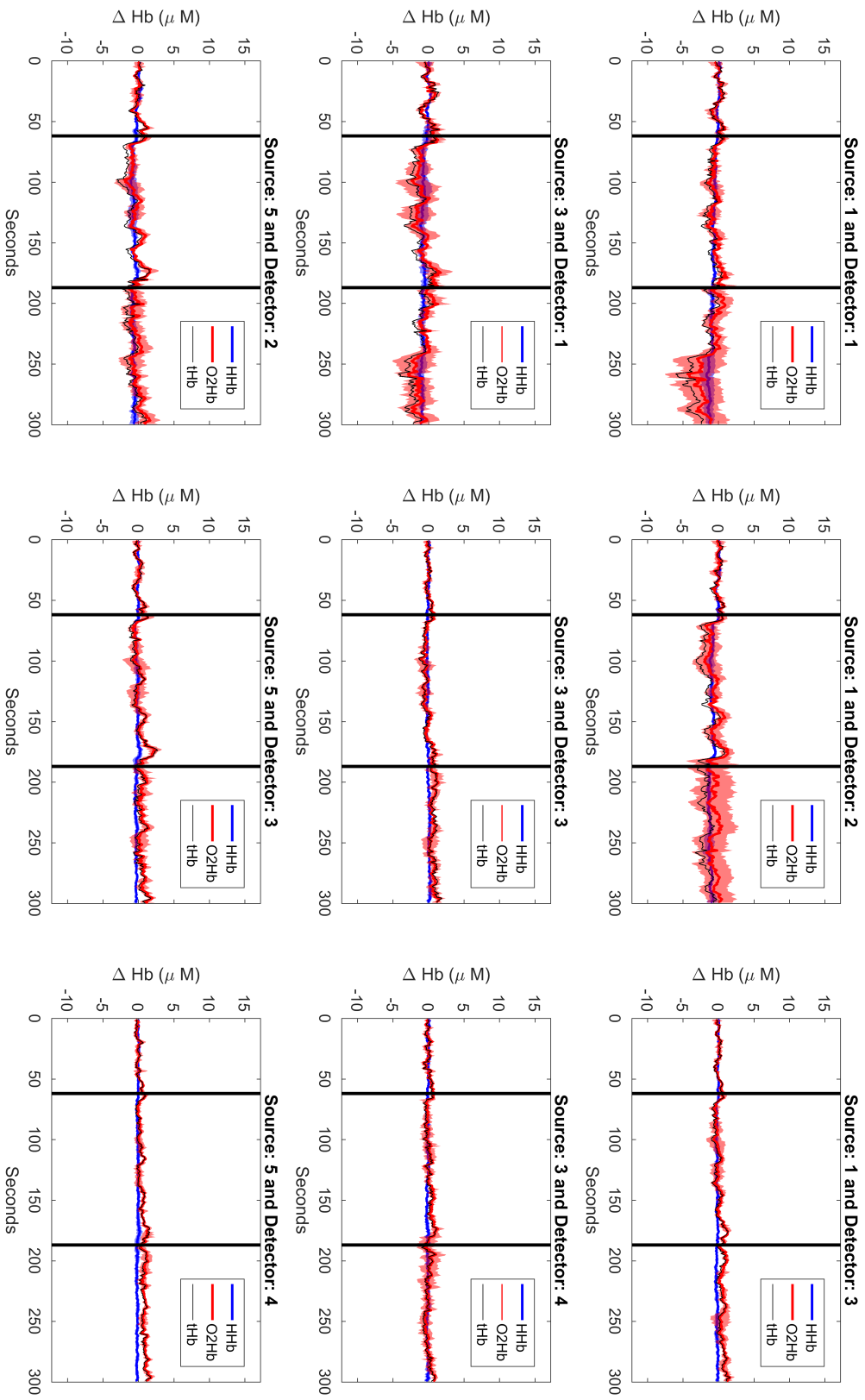


Figure 5.26: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for reference arm during arterial occlusion experiment performed at Subject 2.

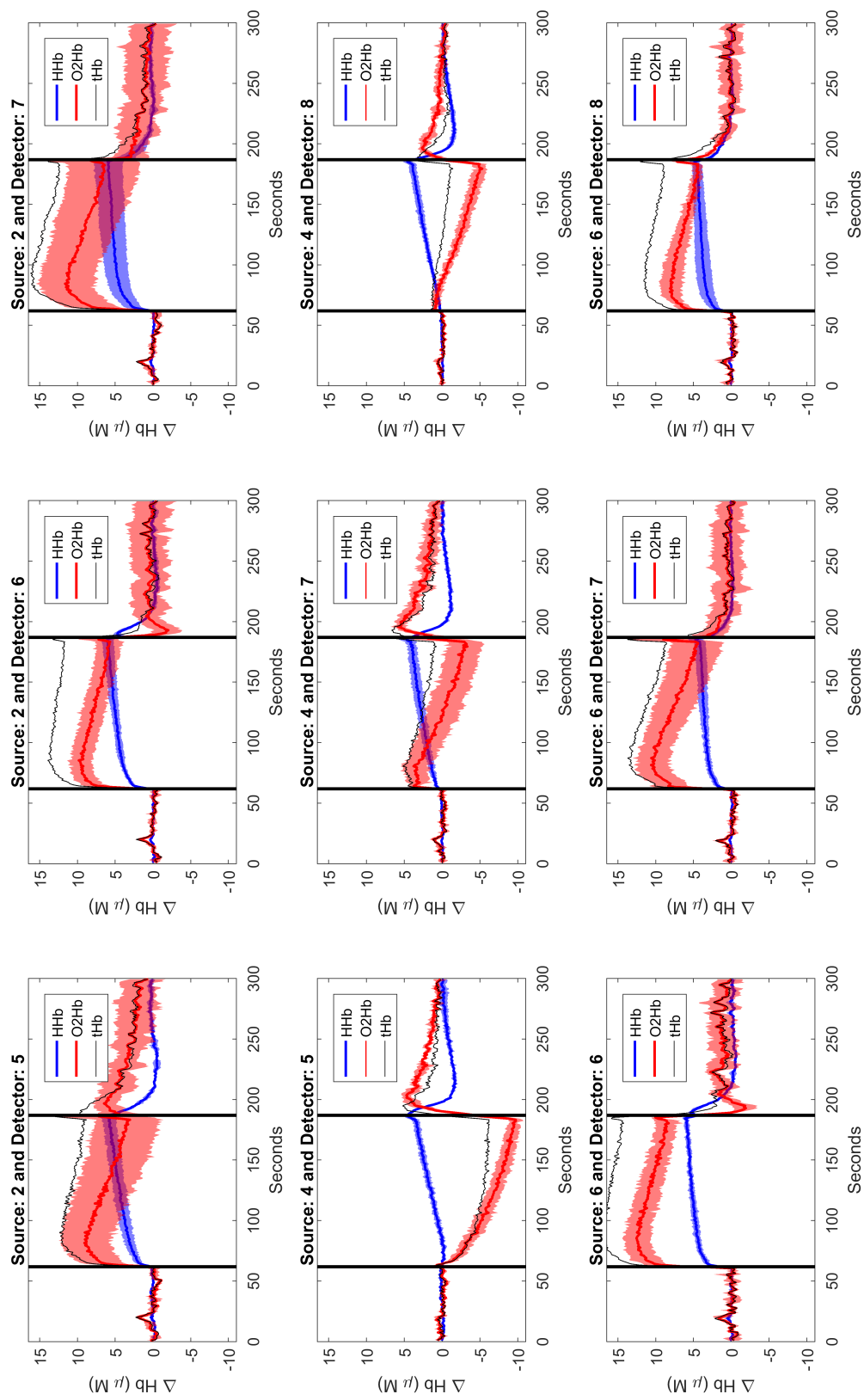


Figure 5.27: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for occluded arm during arterial occlusion experiment performed at Subject 3.

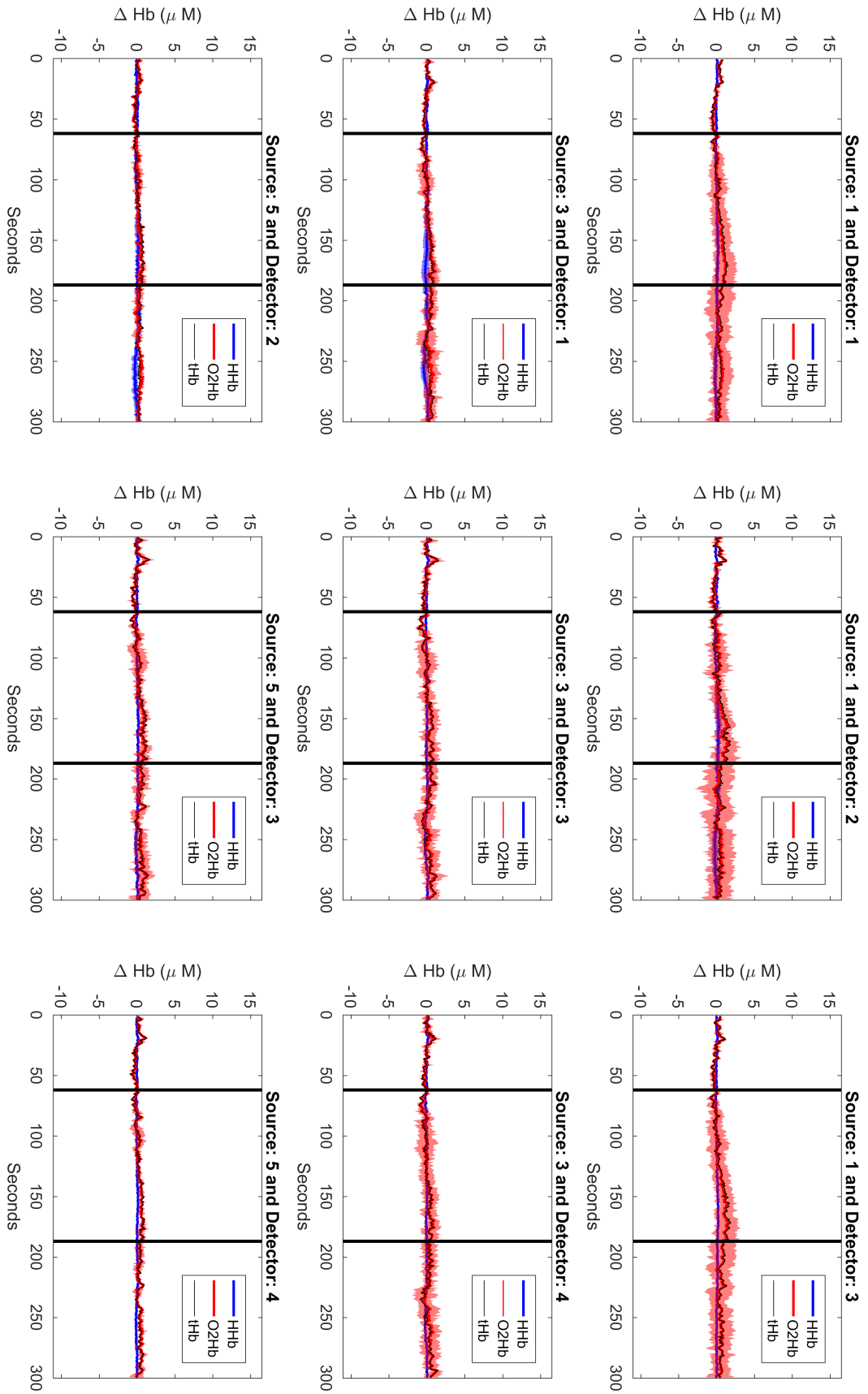


Figure 5.28: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for reference arm during arterial occlusion experiment performed at Subject 3.

## 5.4. Motor cortex activation experiments

After the system has been validated in arm occlusion tests, it is time to perform motor cortex activation experiments. The validity of fNIRS techniques for brain activation imaging has been proven in several studies. For example in [76], finger opposition experiments, with tactile and electrical nerve stimulation were reported. In this paper, the authors demonstrated an increase in the concentration of oxygenated hemoglobin and a decrease of deoxygenated hemoglobin in the contralateral hemisphere in agreement with functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (PET) studies. This study also stressed the robustness and reproducibility of the results for different subjects. After, other works have validated fNIRS technique for brain imaging and neuroscience research [90, 105, 29, 103, 73]. Since the beginning, time-resolved diffuse optics technology has also been used to measure cortical activation [150, 182, 211, 197]. Most advanced systems had up to 64 collection points with a minimum acquisition time of 5 ms per channel [43], with capabilities to measure both brain hemispheres at the same time.

Soon after, first brain cortex tomography approaches were reported [96, 97, 80]. In the second decade of the century, some labs started to image the whole human brain using a high density of source-detectors probes with continuous-wave technology [64, 142]. More recently, in [39], a continuous-wave HD-DOT wearable system is presented. However, in my point of view, one of the main weaknesses of some of these tomographic studies is that they do not take advantage of fNIRS temporal resolution.

In this work, I will present series of time-frame tomographic reconstructions, that will show the evolution of motor cortical activation in depth. Up to my knowledge, this is the first time that, using diffuse optics technology, a continuous flow of tomographic reconstructions frames from a cortical activation are shown.

In the next sections, I will describe the motor cortex activation experiments, the data analysis and the performed tomographic reconstructions.

### 5.4.1. Experiment description

For these experiments, the subjects performed the finger tapping sequence described in Figure 5.29. The sequence goes as follows: (1) thumb touches the index finger, (2) thumb touches the ring finger, (3) thumb touches the middle finger and (4) thumb touches the pinkie finger. This sequence is done cyclically during the finger tapping period; it is complex enough so the subject has to be focused on the movement to do it correctly. The subject should do it as fast as possible but maintaining the correct order of the sequence during the whole experiment.

The finger tapping sequence induces changes in oxy/deoxygenated hemoglobin at C3 and C4 positions, see Figure 5.30 (Left). It is known that the left hemisphere of the brain controls the right side of the body, and the right hemisphere controls the left side of the body. Therefore, C3 activation implies right hand movement and C4 activation makes left hand to move.

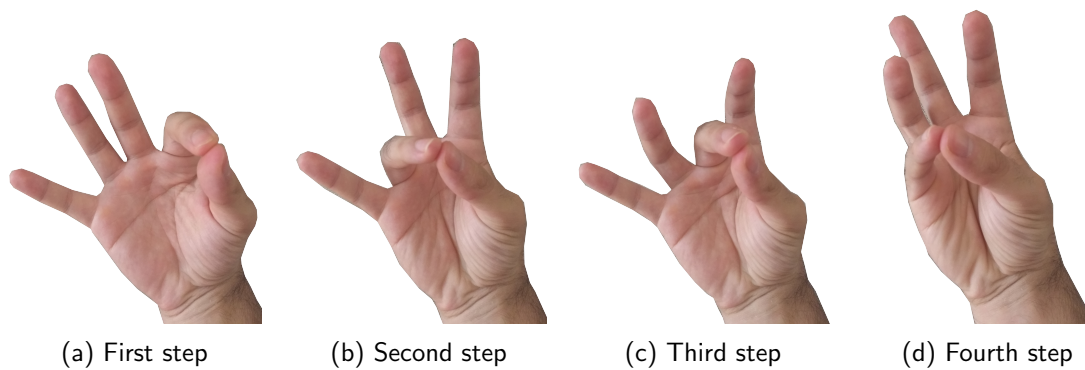


Figure 5.29: Finger tapping protocol to induce changes in oxy/deoxygenated hemoglobin at C3 and C4 motor cortex positions.

The protocol performed by the subjects was divided in three tests. For all the tests, the subject remained lying in a comfortable bed. In the first test, the subject stays without moving for 20 seconds. After, the subject was told to start performing finger tapping with the right hand. 20 seconds after, the subjects was told to stop tapping. The subject stayed for 20 seconds recovering, see Figure 5.31 (Left). The activity was repeated for five times in a row. In the second test, the same protocol was used but the subject had to finger tap with the left hand. In the third test, the subject did not perform finger tapping for 60 seconds, nevertheless after 20 seconds and 40 seconds it was told to start and stop tapping respectively. This test was used as control measurement and was repeated for five times, see Figure 5.31 (Right). Between the three activities, the subjects rested in the bed for ten minutes.

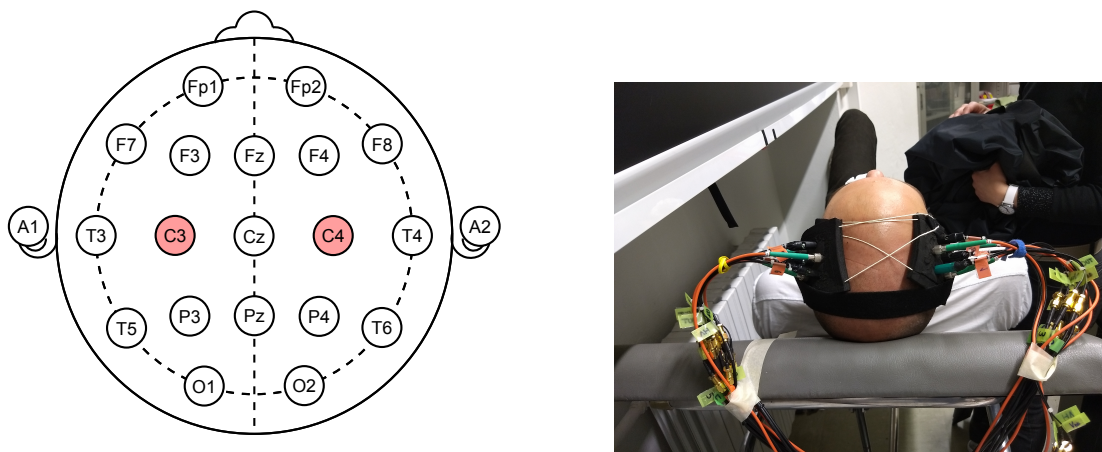


Figure 5.30: (Left) Theoretical C3 and C4 positions in the human head. (Modified figure taken from Wikimedia Commons). (Right) Placement of the probes for a human subject.

A 10–20 system electroencephalography (EEG) adult cap was used to localize C3 and C4 positions. First, the distance from nasion and inion and the length of bitragion coronal arc was measured, see Figure 5.32. Then, the *CZ* position was localized as the intersection of nasion–inion and bitragion coronal arc lines. The probes were located above the C3 and C4 position and slightly shifted to the side to avoid putting the center detector exactly above the cortical positions. This was done to avoid the center blind spot shown at Figure 5.7. In

Figure 5.33 probe location and orientation for each subject is shown.

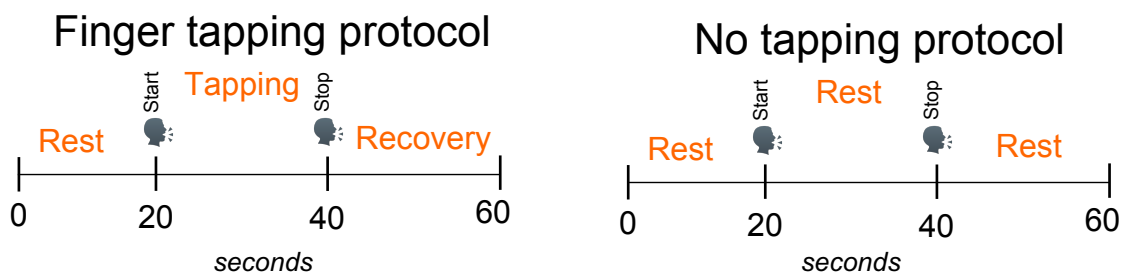


Figure 5.31: (Left) Finger tapping protocol for left and right hand. (Right) Protocol for no finger tapping case. Each of the protocols was repeated for five times to improve the SNR.

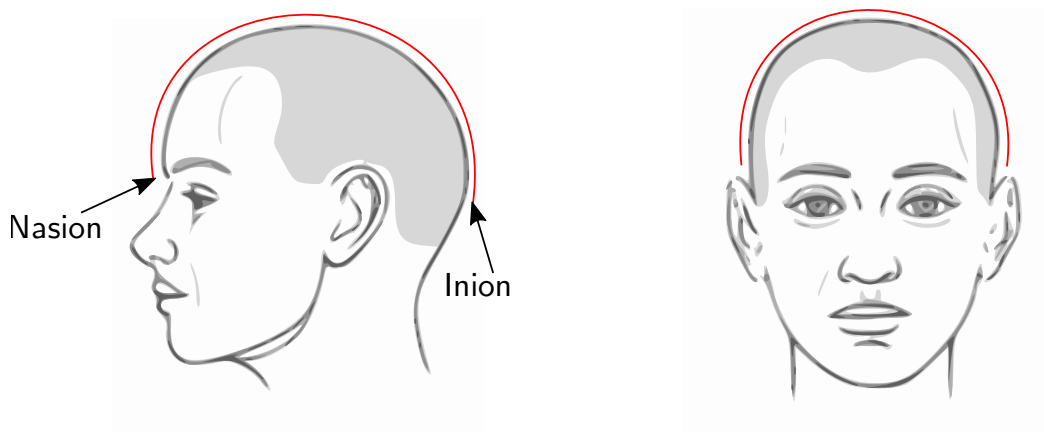


Figure 5.32: (Left) Red line indicates the distance from nasion to inion. (Right) Bitrignon coronal arc.

The expected evoked response from finger tapping should increase the blood flow at the respective motor cortex position. In these responses, both oxygenated hemoglobin and total hemoglobin increase due to the increase of blood flow and volume at the motor cortex region. The physiological changes are related to the neurovascular coupling, that is, the activation of neurons induces blood vessels dilatation around the brain activation area [32] and therefore blood flow and volume is increased. The increase in the arterial blood flow also causes deoxygenated hemoglobin to decrease, this phenomenon is known as *washout effect* [106]. At the same time, an increase in neural activity also induces an increase in the cerebral oxygen metabolism ( $CMRO_2$ ) which converts oxygenated hemoglobin into deoxygenated. However, the increase in blood flow compensates this effect and the net result is that blood oxygen saturation is increased in the activation area [77]. Therefore, the expected result from a motor activation is an increase of oxygenated and a decrease of deoxygenated hemoglobin.

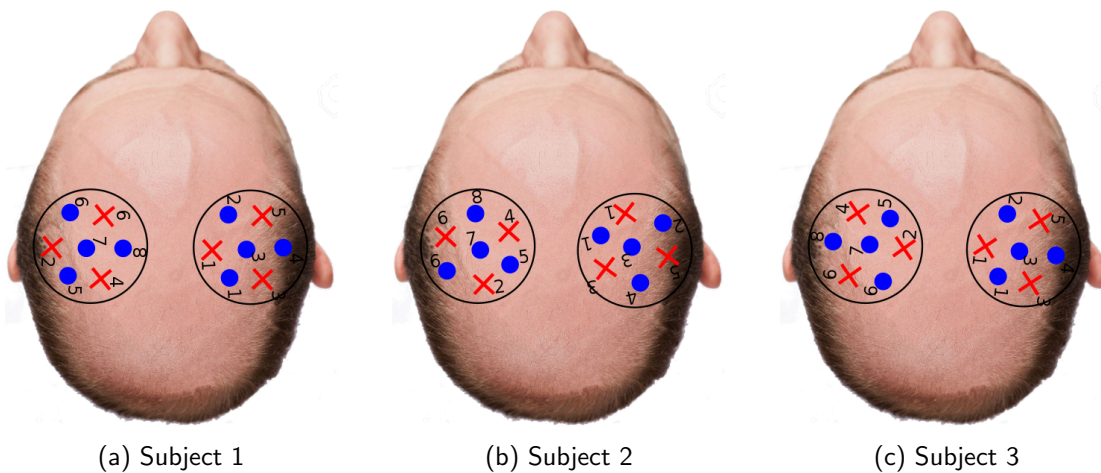


Figure 5.33: Probes locations for each of the subjects. Each probe has four detectors (blue dots) and three light fibers (red dots).

#### 5.4.2. Results and discussion

One of the critical aspects to measure brain activation in human heads is to have a good count rate at most detectors. For this reason, the selected volunteers (three male) had naturally few hair in the head and it was shaved before performing the experiments. In Figure 5.34, the average count rate over the three protocols for each wavelength and subject is shown. At first glance, it is evident that Subject 1 has a larger count rate for most channels in comparison with rest of subjects. The other subjects only have a good enough count rate for few channels.

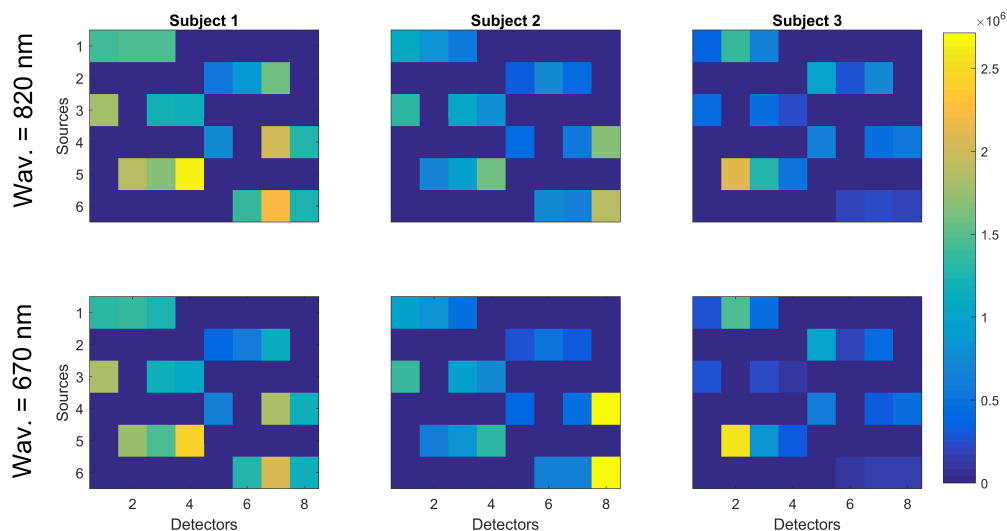


Figure 5.34: Average count rate over the three finger tapping cases. Each column represents one subject and first and second row represent 820 nm and 670 nm wavelengths respectively.

For motor cortex experiments, the time series results were obtained using time-of-flight analysis explained at Subsection 5.2.2. In Figure 5.35, the time series results for Subject 1

during left hand tapping experiment at right hemisphere are shown. For channels 1–1, 1–2, 1–3, 5–2, 5–3 and 5–4 it is clearly seen the increase of oxygenated hemoglobin during finger tapping and the expected decrease of deoxygenated hemoglobin. Moreover, the total hemoglobin also increases during activation as expected. It can also be seen a delay of five to seven seconds between the start of the finger tapping and the maximum changes of oxy/deoxygenated hemoglobin, this delay is in accordance with the results published in [33, 144]. Quantification levels of hemoglobin concentration changes are also very similar to results published in [220, 39]. As expected, no activation effect is seen at left hemisphere, see Figure 5.36. For right hand tapping experiment, activation is also clearly appreciated at channels 4–7, 4–8, 6–7 and 6–8, see Figure 5.37. In this case, the values of hemoglobin differences are lower compared to left hand tapping. Asymmetries in brain morphology, handedness and activation differences between the two hemispheres during motor-related tasks are a wide field of research. The results for this particular (right handed) subject could suggest that more activation is seen at right hemisphere because left (non dominant) hand movements require more concentration. Nevertheless, in several studies [117, 51, 214], evidence was given that left motor cortex is activated substantially for right handed subjects in comparison with right hemisphere. No activation was measured at right hemisphere as expected, see Figure 5.38. The results for no hand tapping were included in Appendix H to avoid over-saturating the main text – as expected, no activation is appreciated.

For Subject 2, some activation is seen for left hand tapping in the right hemisphere, see channels 1–1, 1–3 and 3–1 at Figure 5.39. However, the activation values are lower compared to Subject 1. A possible explanation is that the count rate for this subject was lower than for Subject 1, see Figure 5.34. As expected, no activation is seen for left hemisphere when left hand tapping is performed. For right hand tapping no activation is seen, see Figures H.3 and H.4 – a possible explanation is that the count rate for detectors at left hemisphere were very low (except for detector 8 which, by the way, was furthest from C3 position). No activation is seen for no hand tapping protocol, see Figures H.5 and H.6.

For Subject 3 activation is not clearly seen in any of the protocols. At channel 1–1 from right hemisphere at left hand tapping (Figure H.8) some activation may be guessed but the significance is not sufficient enough. The results for this subject are understandable taking into account the very low count rate that it had. Time series of Subject 3 for during the rest of protocols can be seen at Appendix H.

The tomographic reconstructions for motor cortex activation experiments can be found at [74]. As with arm occlusion experiments, the reconstructions for each time frame were averaged over the five repetitions.



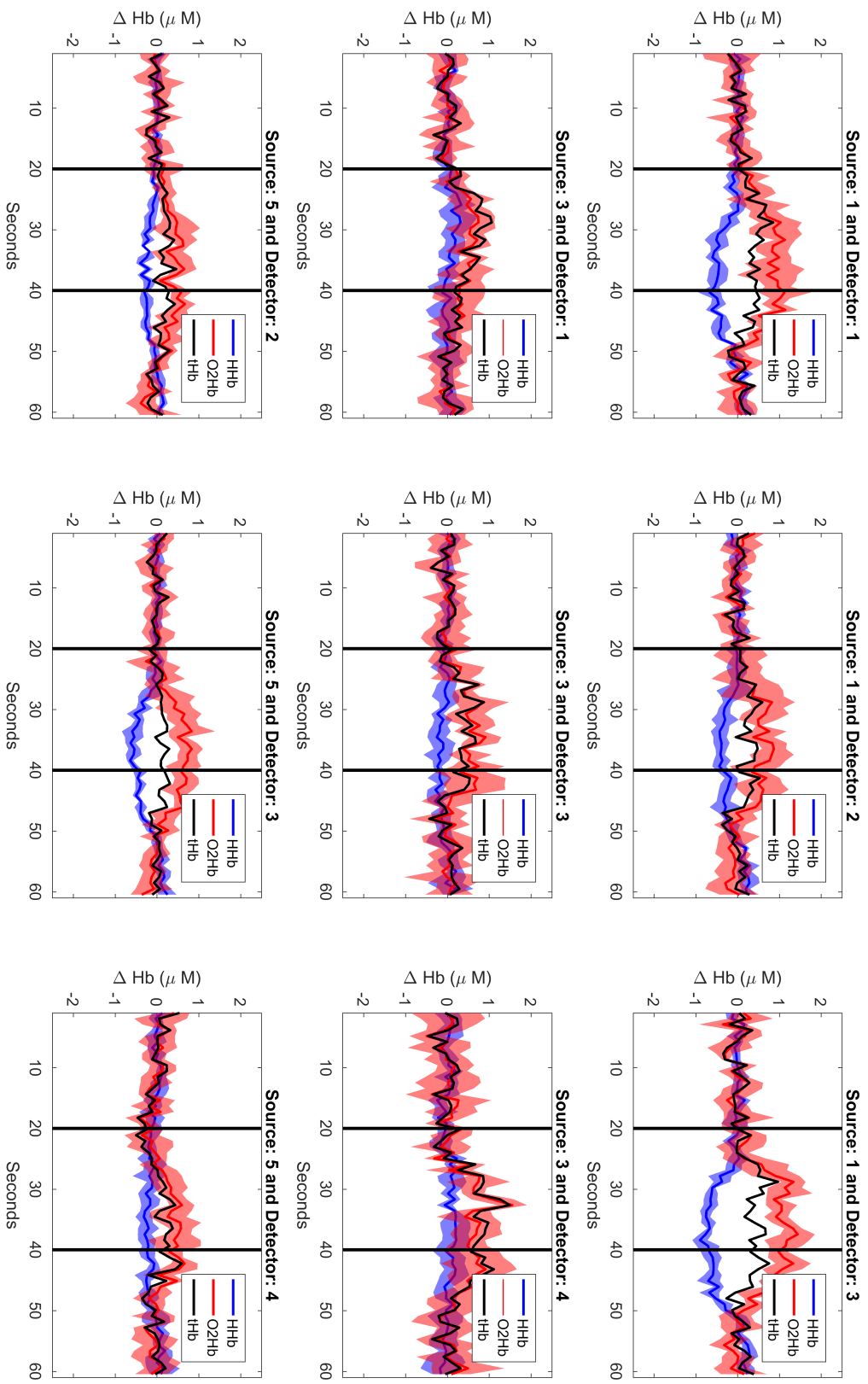


Figure 5.35: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for right hemisphere during left hand tapping performed by Subject 1. Black time series indicates the total hemoglobin value. The shadows indicate the standard deviation over five repetitions. Finger tapping happened between black vertical lines.

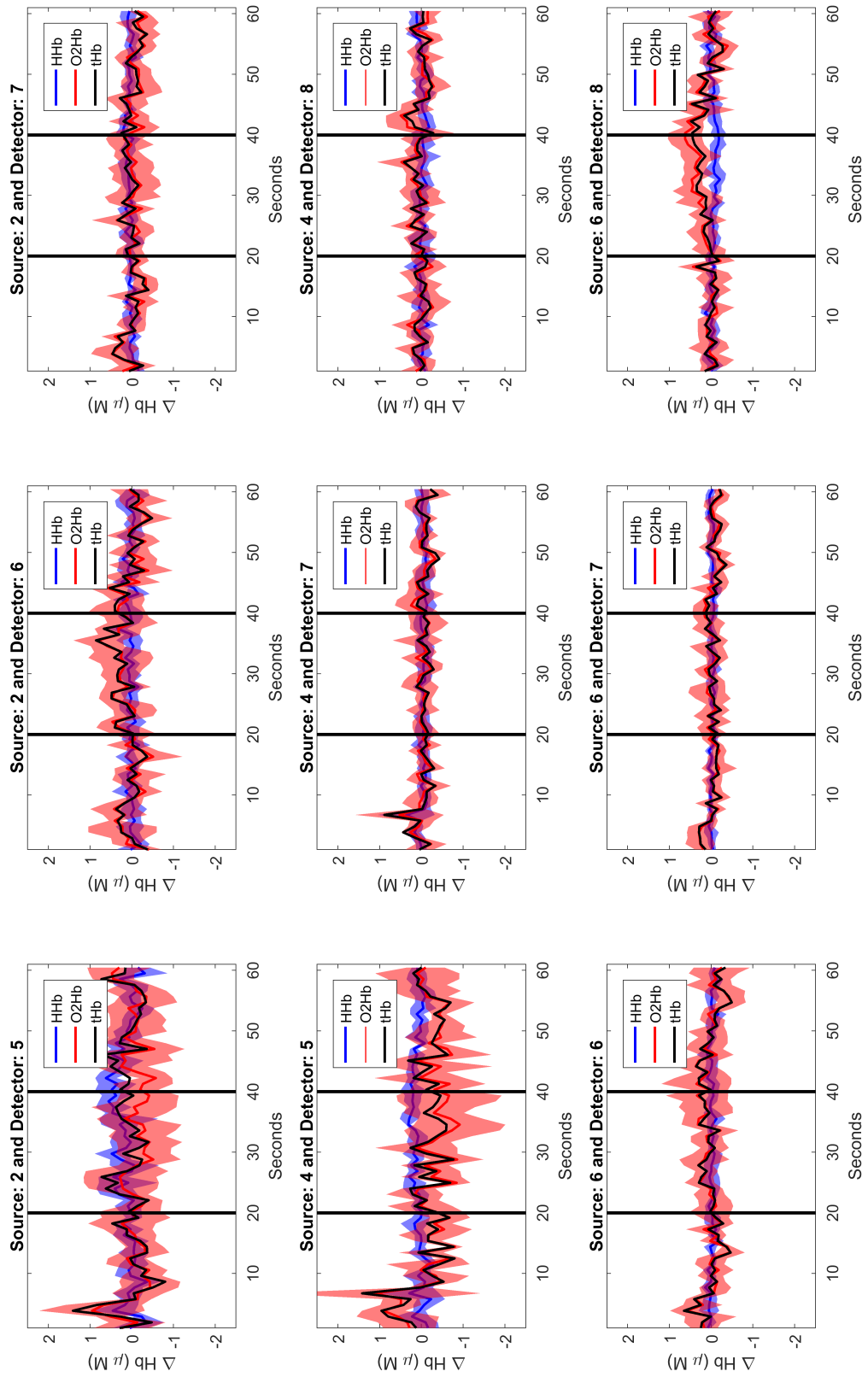


Figure 5.36: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for left hemisphere during left hand tapping performed by Subject 1.

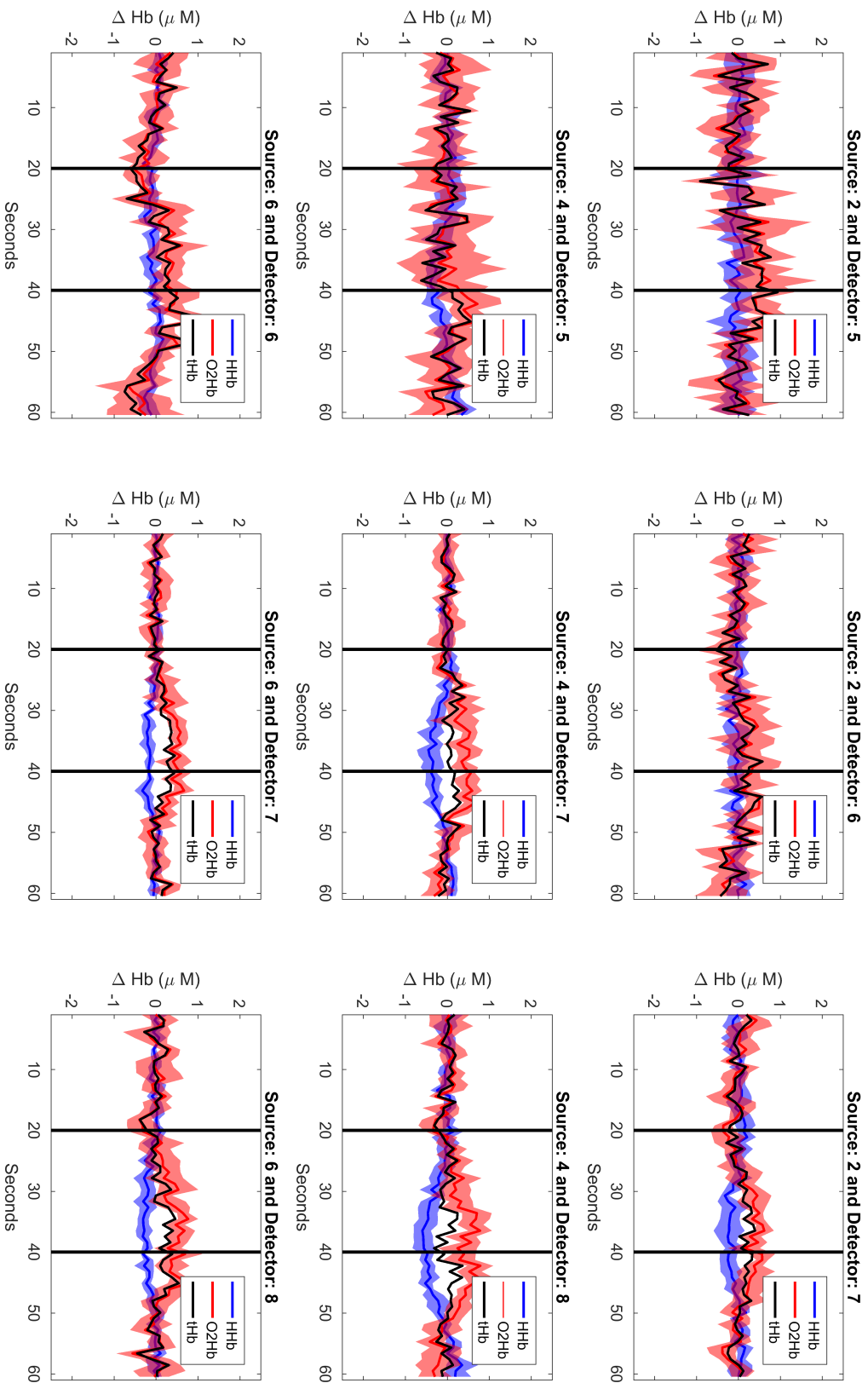


Figure 5.37: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for left hemisphere during right hand tapping performed by Subject 1.

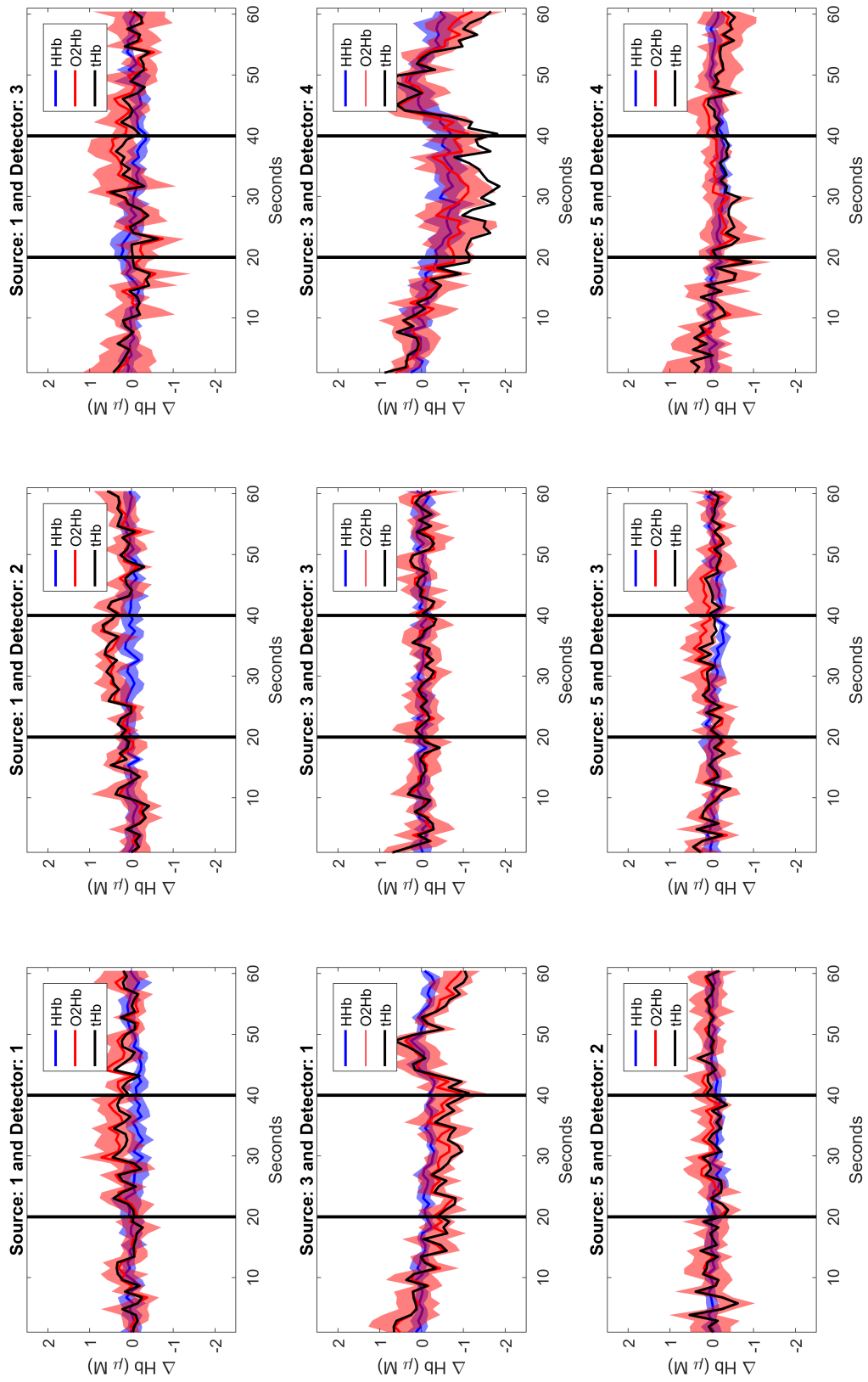


Figure 5.38: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for right hemisphere during right hand tapping performed by Subject 1.

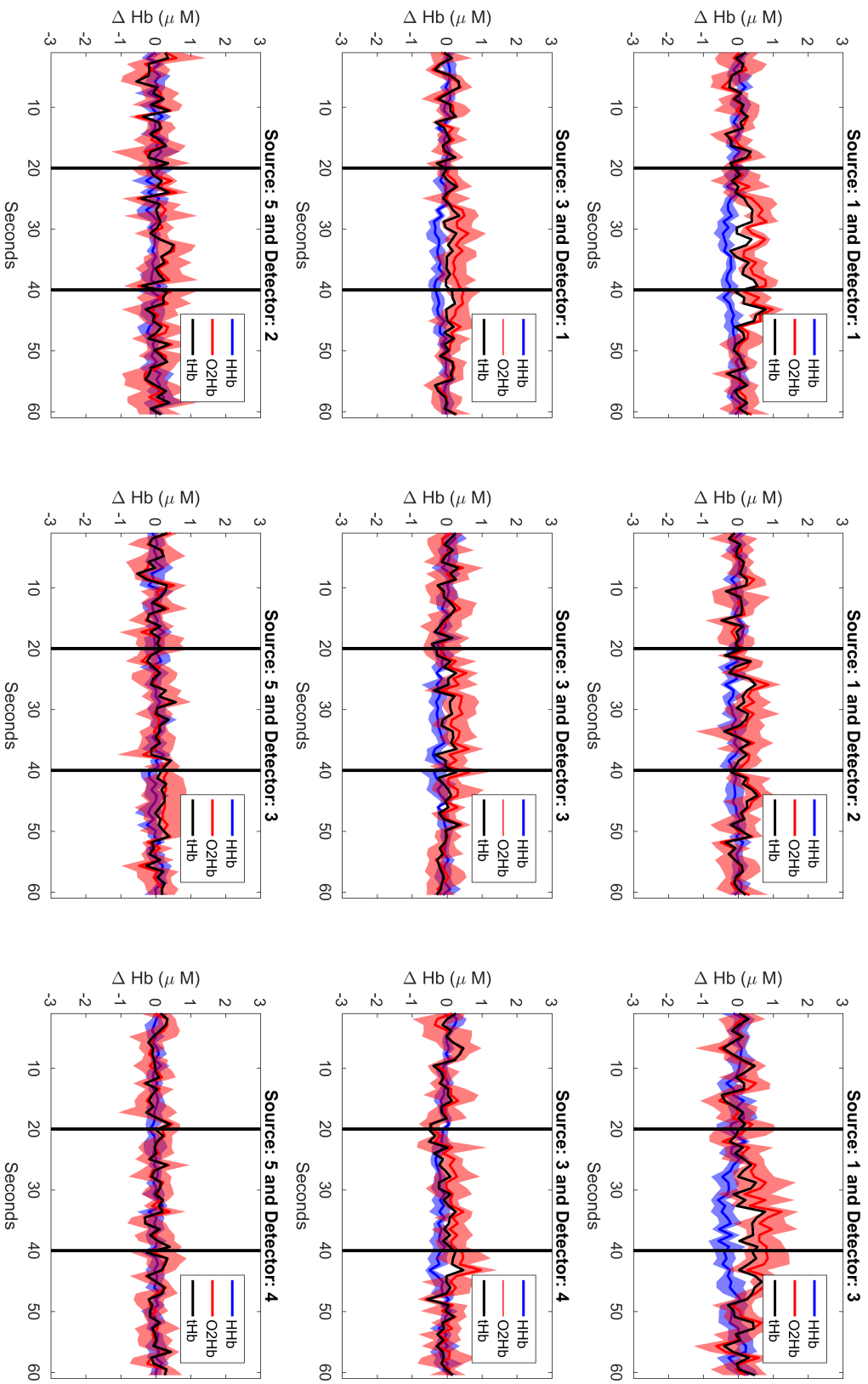


Figure 5.39: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for right hemisphere during left hand tapping performed by Subject 2.

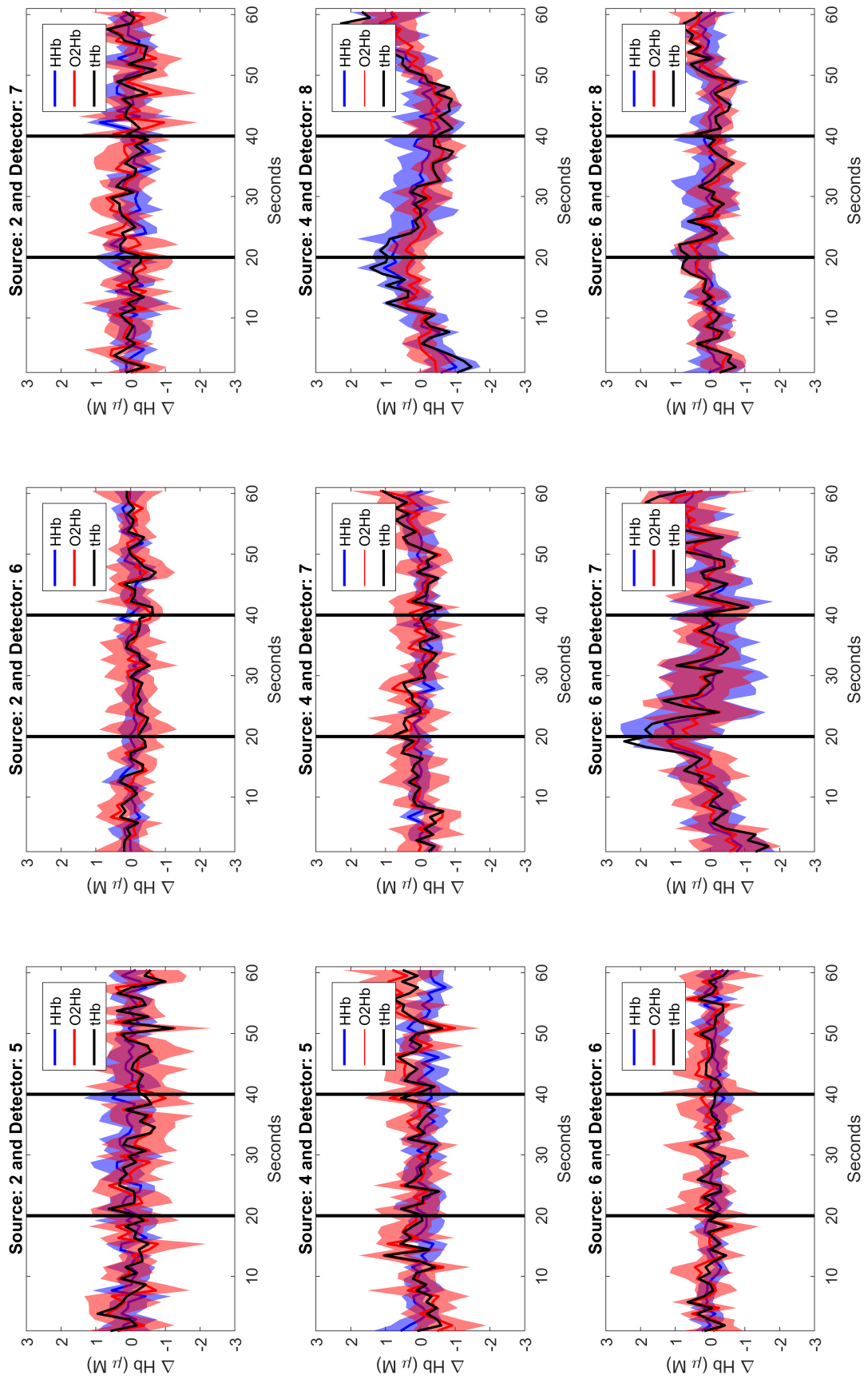


Figure 5.40: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for left hemisphere during left hand tapping performed by Subject 2.

For Subject 1, activation can be clearly seen for left and right hand tapping (please, go to `motor_cortex_experiments` and see `S1_Hb_left-tap.avi` and `S1_Hb_right-tap.avi` videos to observe both hemispheres at the same time, see videos `S1_Hb_left-tap_right-hemi_thres.avi` and `S1_Hb_right-tap_left-hemi_thres.avi` which were thresholded). In Figure 5.41, activation during left hand tapping is shown. Increase of  $O_2Hb$  and decrease of  $HHb$  is clearly seen. The position of the activation is more biased to the center of the brain rather than the boundary. To test if the activation is localized at the same position during all repetitions, in folder `motor_cortex_experiments/repeatability`, I included some videos showing the increase of  $O_2Hb$  and decrease of  $HHb$  concentration for each of the repetitions performed by Subject 1 (particularly interesting are `S1_O2Hb_hemi-right_left-tap.avi`, `S1_HHb_hemi-right_left-tap.avi`). At right hemisphere for left hand tapping the decrease of  $HHb$  concentration is at the same location during the five repetitions, see Figure 5.42. At the same time,  $O_2Hb$  concentration is also increasing for all repetitions, see Figure 5.43. The activation is located between 1.5 to 2 cm depth which is the expected activation depth for an adult subject [105].

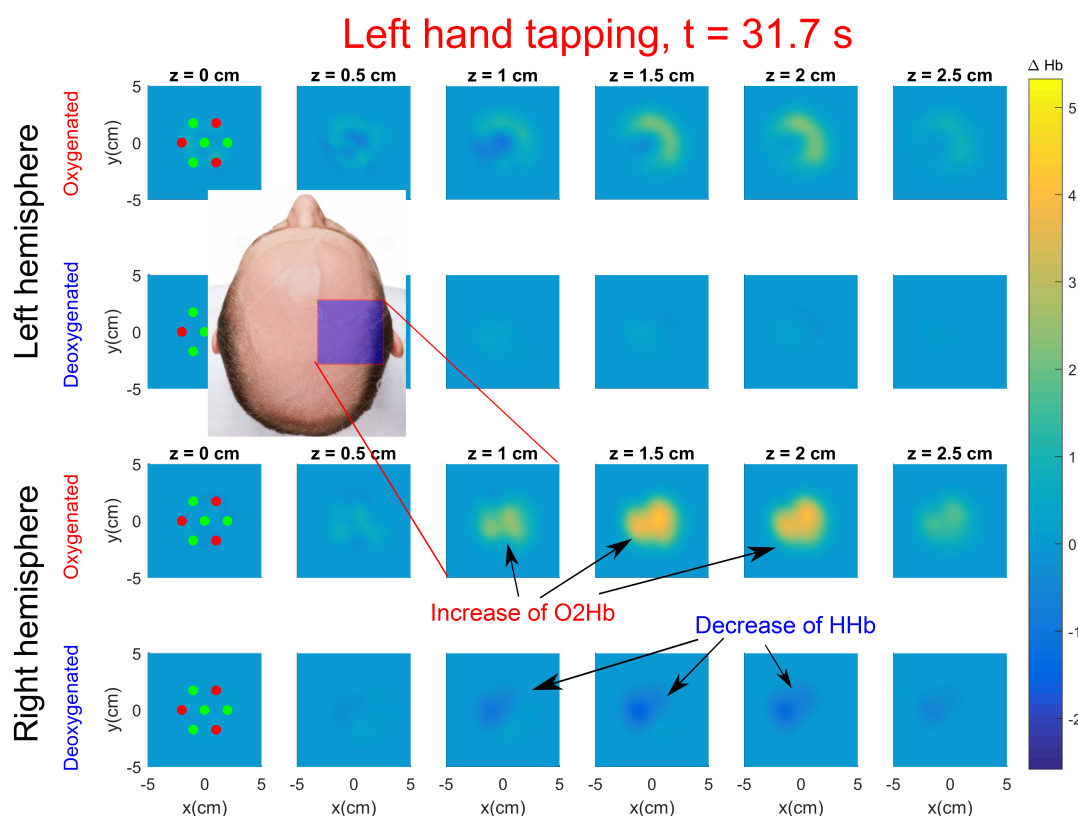


Figure 5.41: Tomographic reconstruction of activation at right hemisphere during left hand tapping experiment performed by Subject 1.

For right hand tapping, see Figure 5.44, activation is also clearly seen at left hemisphere (see video `S1_Hb_right-tap.avi`). The activation position is more shifted to the center of the brain and constant for most repetitions, see Figures 5.46 and 5.45 (see videos `S1_O2Hb_hemi-left_right-tap.avi`, `S1_HHb_hemi-left_right-tap.avi`). Depth values are similar to left hand tapping. An increase of oxygenated blood is also seen at ipsilateral



(right) hemisphere although there is no decrease of deoxygenated blood – for left hand tapping this effect could also be slightly seen. A similar phenomenon was also reported in [141]; the authors discuss that complex hand movements with dominant hand or simple hand movements with non-dominant hand tend to require more cortical activity and they also recruit some activity from ipsilateral hemisphere. Similar conclusions are reached at [219], although they suggest that these differences disappear when subjects are overtrained.

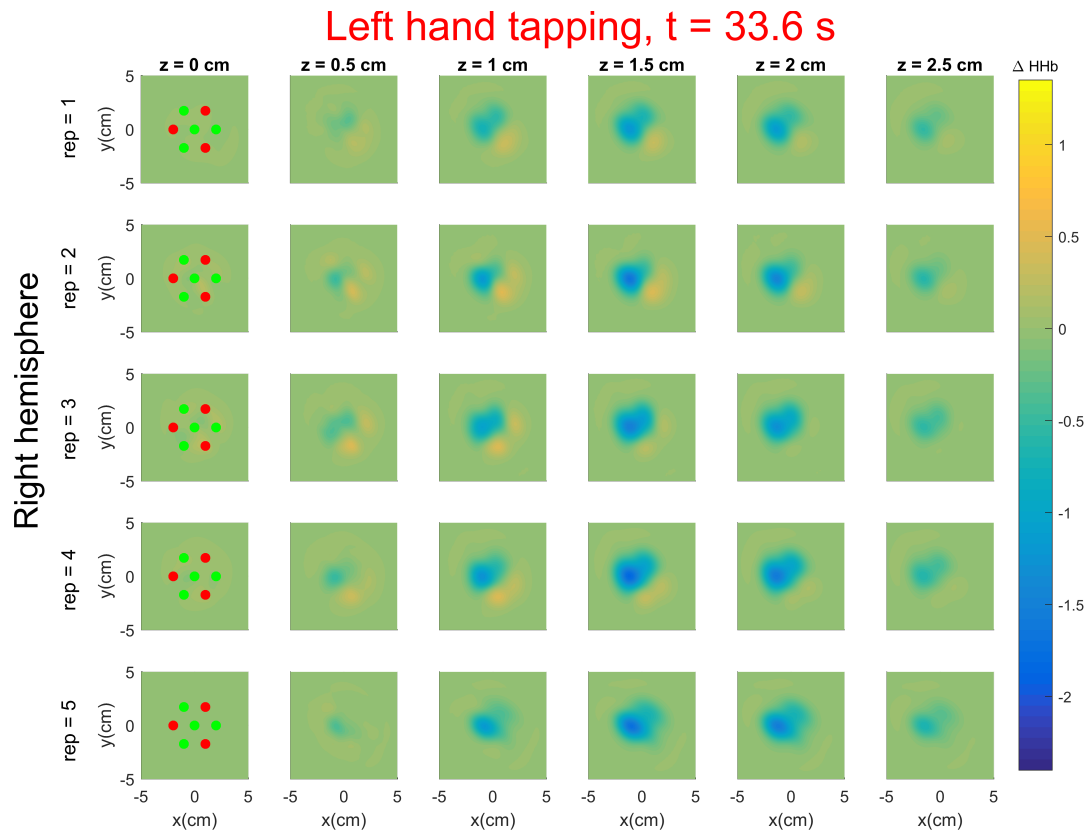


Figure 5.42: Tomographic reconstruction at  $t = 33.6$  s of HHb concentration at right hemisphere for left hand tapping experiment during the five repetitions performed by Subject 1.

An important difference between time-series data analysis and tomographic reconstructions is that for the latter the hemoglobin concentration changes values are very similar between right and left-hand tapping (note that for the time-series analysis the activation was larger for left-hand tapping). One interpretation is that the spatial and in-depth information obtained from tomography manages to balance better the changes in hemoglobin concentration, resulting in better estimations. Nevertheless, further experiments have to be done to confirm this hypothesis.



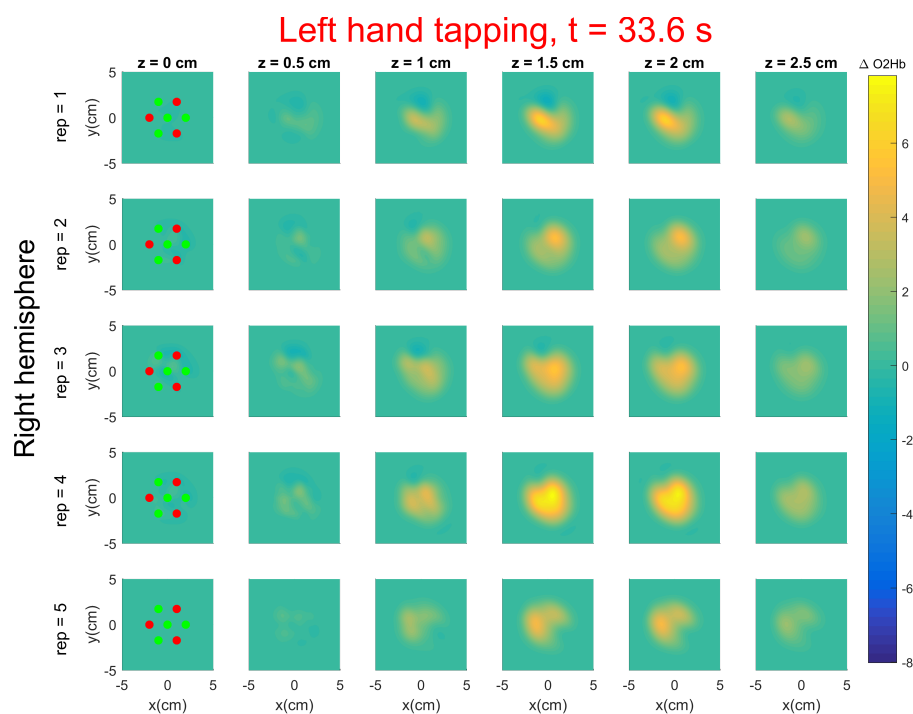


Figure 5.43: Tomographic reconstruction at  $t = 33.6$  s of  $O_2Hb$  concentration at right hemisphere for left hand tapping experiment during the five repetitions performed by Subject 1.

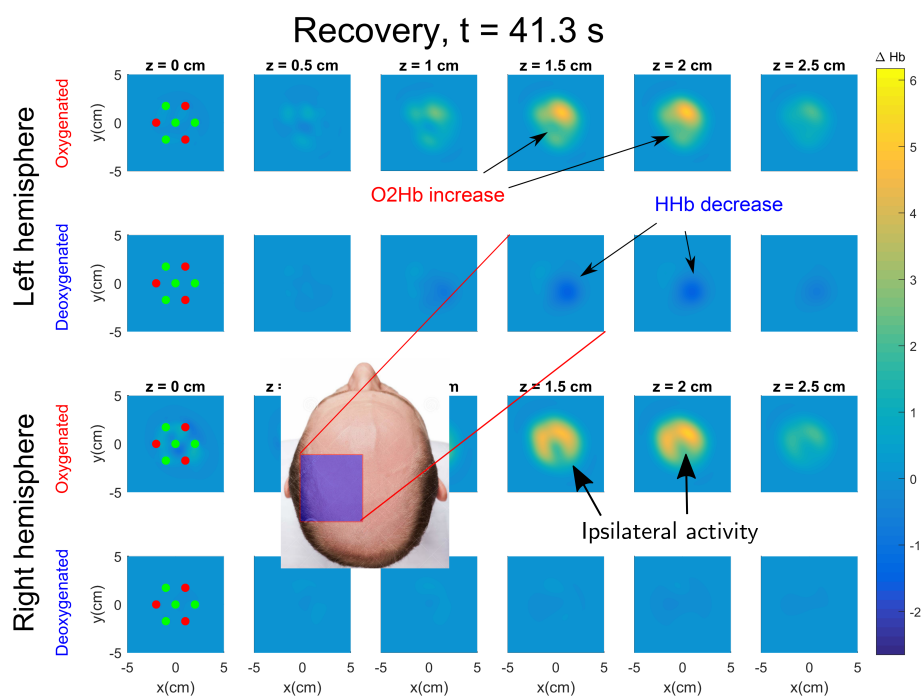


Figure 5.44: Tomographic reconstruction of activation at left hemisphere during right hand tapping experiment performed by Subject 1.

Videos of Subject 2 can also be found in the same repository. From the averaged reconstructions no activation could be seen. The reconstructions for each repetition with left hand tapping at right hemisphere are interesting (videos S2\_O2Hb\_hemi-right\_left-tap.avi and S2\_HHb\_hemi-right\_left-tap.avi). The decrease pattern of HHb at right hemisphere is repeated several times (see Figure 5.47) although the concentration difference is very low (around  $\Delta\text{HHb} = -0.6$ ) and therefore is difficult to appreciate in the averaged results. For this subject, the decrease of HHb is located shallower than for first subject, maximums are between 1 to 1.5 cm depth.

For third subject, similar results as for Subject 2 are obtained. The reconstructions averaged over all repetitions (see LC\_Hb\_left-tap.avi and LC\_Hb\_right-tap.avi) do not show a decrease of HHb concentration, although there is a significant increase of O<sub>2</sub>Hb in the contralateral hemisphere for left hand tapping and ipsilateral hemisphere for right hand tapping, that is, for right hemisphere (where best count rate for Subject 3 is obtained, see Figure 5.34). However, when the results for each repetitions are shown (interesting are results for right hemisphere are shown at LC\_HHb\_hemi-right\_left-tap.avi, LC\_HHb\_hemi-right\_right-tap.avi) a common decrease of HHb during finger tapping is clearly seen, see Figure 5.48 and 5.50. Although, the decrease is one magnitude order lower than O<sub>2</sub>Hb increase (see Figure 5.49 and 5.51), the repeatability for all repetitions is significant. The decrease of HHb is located at similar depth to Subject 2, maximums are between 1 to 1.5 cm depth.

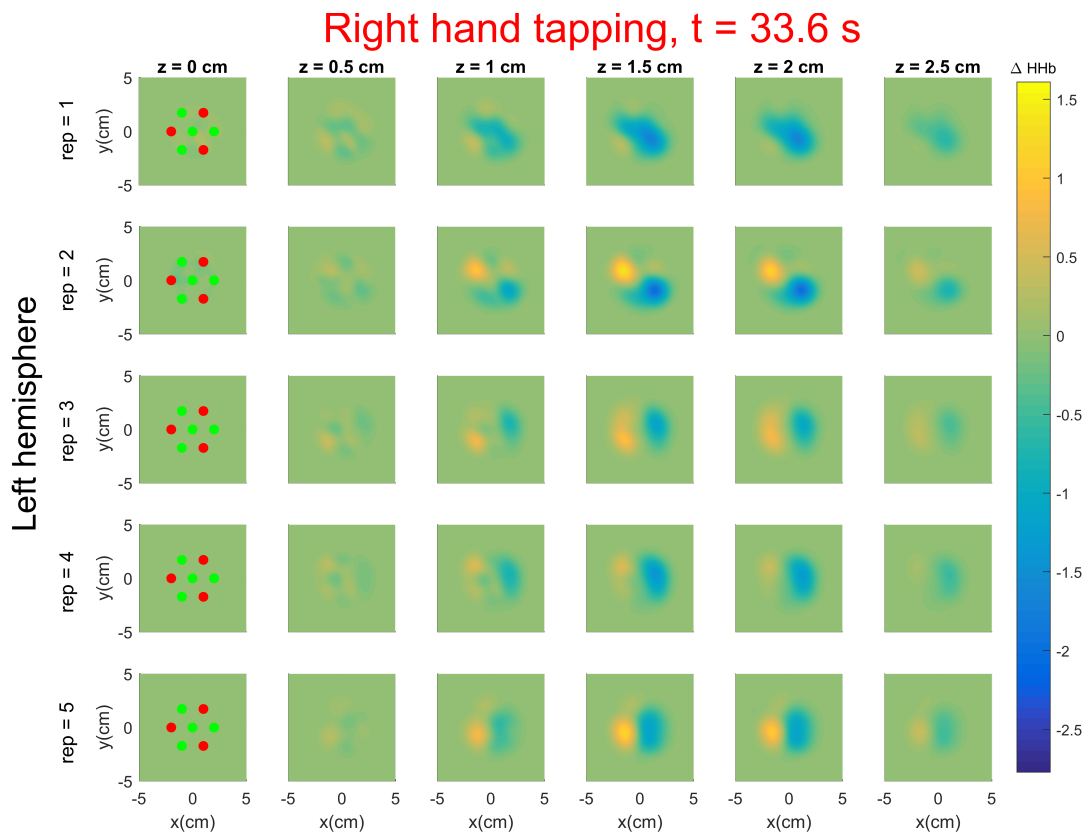


Figure 5.45: Tomographic reconstruction at  $t = 33.6$  s of HHb concentration at left hemisphere for right hand tapping experiment during the five repetitions performed by Subject 1.

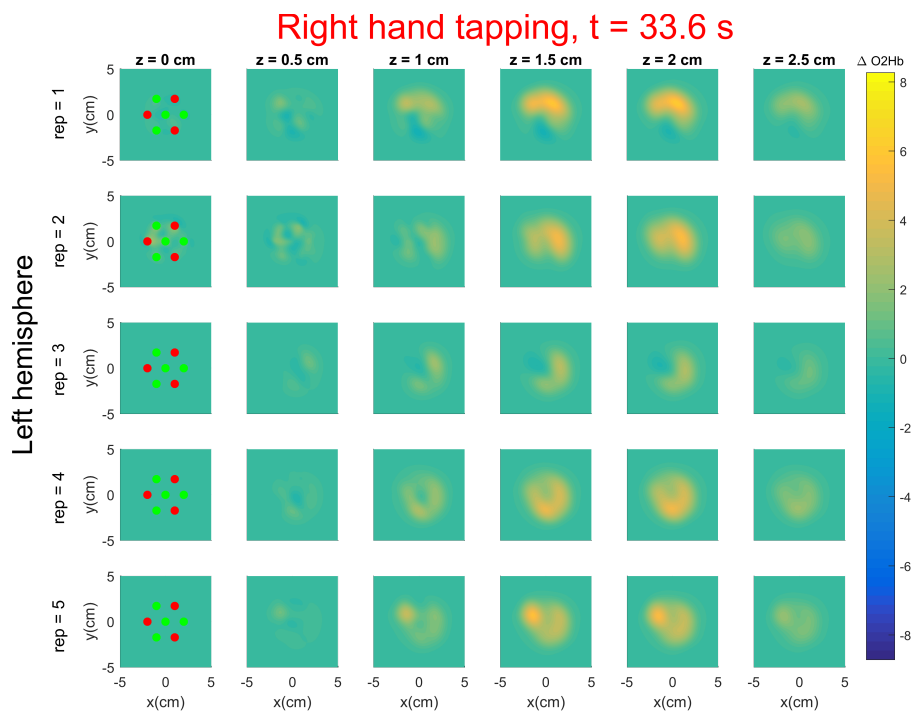


Figure 5.46: Tomographic reconstruction at  $t = 33.6$  s of  $O_2Hb$  concentration at left hemisphere for right hand tapping experiment during the five repetitions performed by Subject 1.

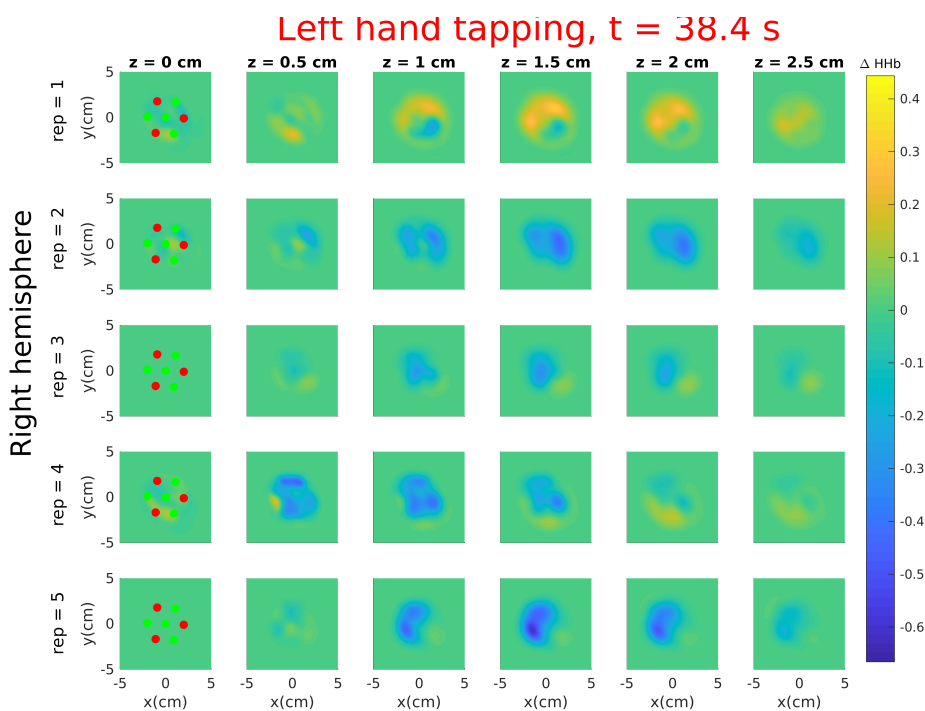


Figure 5.47: Tomographic reconstruction at  $t = 38.4$  s of  $HHb$  concentration at right hemisphere for left hand tapping experiment during the five repetitions performed by Subject 2.

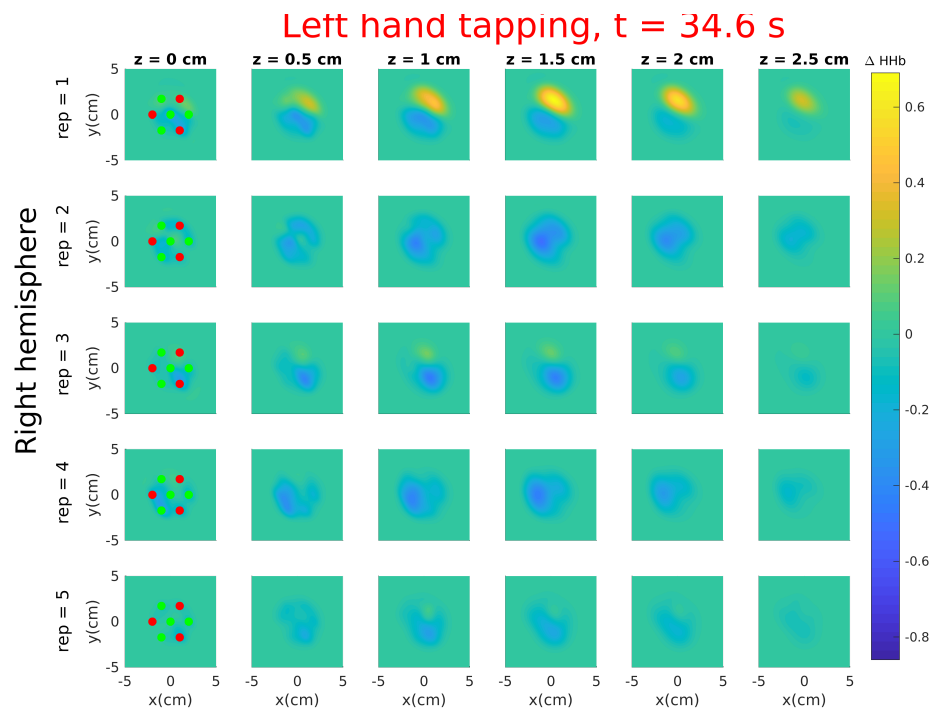


Figure 5.48: Tomographic reconstruction at  $t = 34.6$  s of HHb concentration at right hemisphere for left hand tapping experiment during the five repetitions performed by Subject 3.

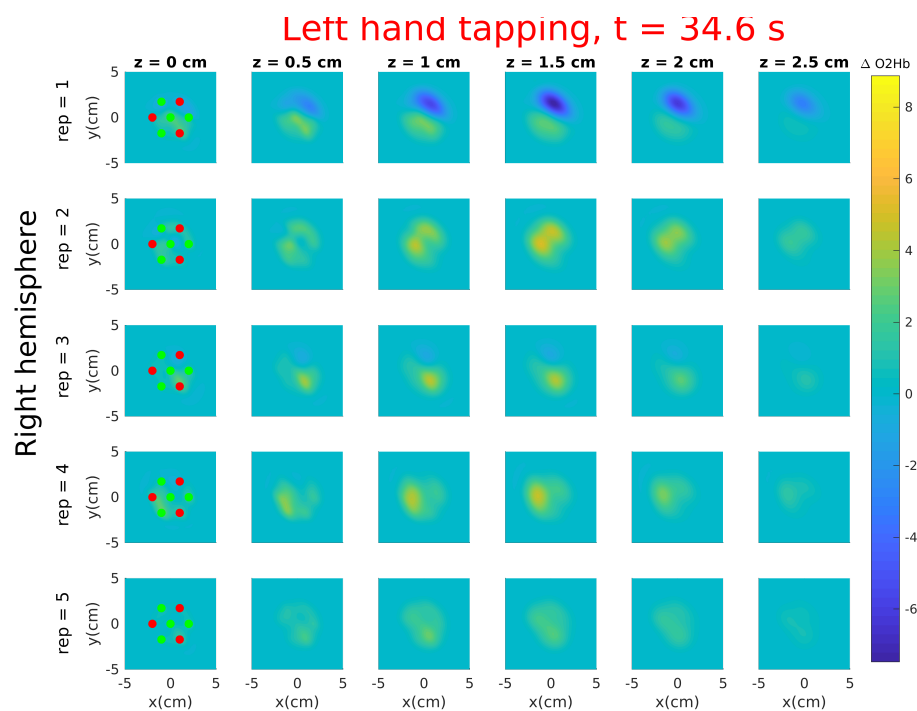


Figure 5.49: Tomographic reconstruction at  $t = 34.6$  s of O<sub>2</sub>Hb concentration at right hemisphere for left hand tapping experiment during the five repetitions performed by Subject 3.

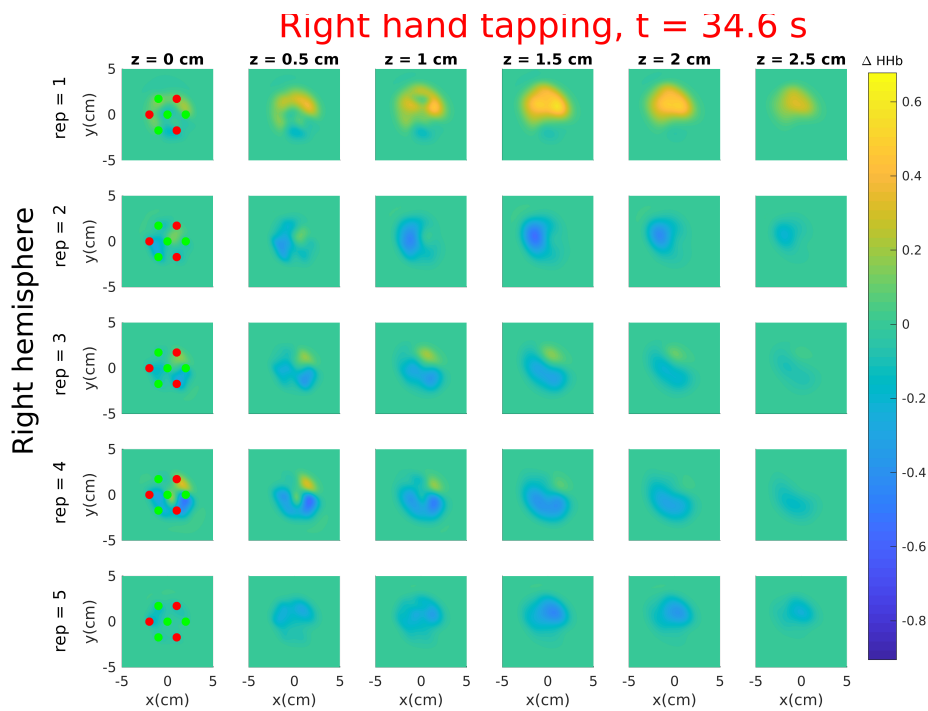


Figure 5.50: Tomographic reconstruction at  $t = 34.6$  s of HHb concentration at right hemisphere for right hand tapping experiment during the five repetitions performed by Subject 3.

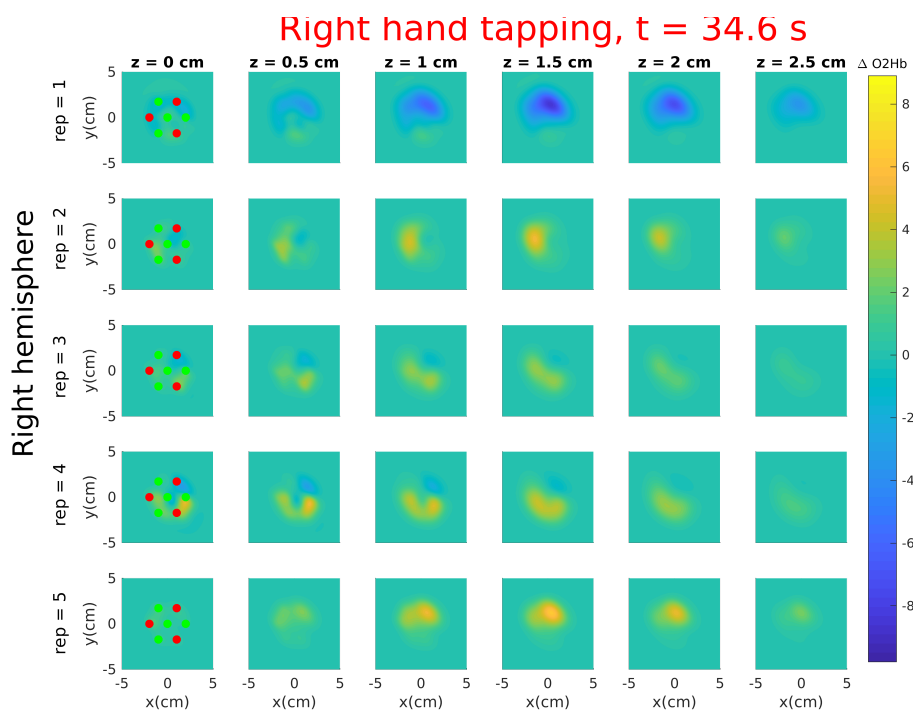



Figure 5.51: Tomographic reconstruction at  $t = 34.6$  s of  $O_2Hb$  concentration at right hemisphere for right hand tapping experiment during the five repetitions performed by Subject 3.

## 5.5. Conclusions

A fast-switching time-resolved diffuse optical system was validated with arm occlusion and motor cortex activation experiments. The data analysis and tomographic reconstruction performed using arm occlusion measurements validated the performance of the system. After, the lab prototype was also tested in motor cortex activation experiments. Due to the simple probe attachment system that was used, only for one out of three subjects the count rate was good enough to measure cortex activation consistently. For this subject, motor cortex activation at contralateral hemisphere was clearly seen for left and right hand tapping. I performed tomographic reconstructions for all the measurements obtained during the experiment, from which I obtained 63 frames (0.96 s duration each) of tomographic reconstructions from which the evolution of motor cortex activation is seen. The results show that the activation position is constant during the experiment, both spatially and in depth. Up to my knowledge, this is the first time that, using diffuse optics technology in combination with tomographic algorithms the evolution of brain activation is shown during a whole experiment.



## The BitMap open dataset on performance assessment of diffuse optics instruments

 OPEN SCIENCE movement has increased in popularity during the last years. Its goal is to make scientific research open to all the society by openly publishing not just the results but also the data and software used during research. Open source software has already been quite popular since the appearance of software development on-line hosts; for example, in diffuse optics community there exist several open source software such as NIRFAST [54], NeuroDOT [65] or TOAST [194]. Moreover, to publish the acquired experimental data is growing in popularity between the research community. In some scientific fields like particle physics, Open Data approach is already well established within the community, where in some cases more than one petabyte of well-documented information is freely available [34]. However, there is still a lot of work to do since data publication is not in the standard workflow of many research labs yet. For example, in the diffuse optical spectroscopy and imaging communities, the Open Data approach is still not fully developed. Data from experiments are scattered across the Internet, difficult to compare due to the lack of metadata/documented information or simply not available. Moreover, the lack of public data on standard phantoms also makes standardization of diffuse optics instruments much harder to accomplish since it is difficult to assess and compare the performance of diffuse optics instruments at labs and hospitals. That is, a repository with documented diffuse optical data following well-defined standards is still missing.

Here, I present the dataset obtained from a performance assessment campaign where several European diffuse optics instruments were involved [123]. In this campaign, PhD student from BITMAP-ITN network moved to these laboratories and performed three European protocols using a consistent set of phantoms. The obtained dataset will be organized and uploaded to an Open Data repository in companion with detailed metadata and documentation. A standardized data format, promoted by the Society for Functional Near Infrared Spectroscopy [200], was used to organize the measurements, instrumentation technical information and metadata. This dataset will contain measurements from the largest set of diffuse optics instruments published until now.

In this chapter, I give an introduction to FAIR research data approach and the motivations that made European Commission to encourage its use. After, I describe the measurement



campaign that was carried within the BitMap partners and how I organise those measurements using SNIRF format. Finally, I give some brief conclusions about the project.

## 6.1. Open Data management and FAIR principles

Recently, the European Commission (EC) has indicated that up to 10.2 billion euros per year could be saved by using FAIR research data [218]. The report takes into account not only the cost due to time wasted performing similar experiments but also the wasted money due to cost storage of redundant copies, license prices, research retraction, double funding, lack of interdisciplinary research and loss in potential economic growth. To promote the share of scientific data with the society, the EC have given some recommendations and business models so laboratories can implement FAIR data approaches and sustain them [217]. The EC encourages to make data open to different communities but at the same time it acknowledges that in some projects is not feasible due to intellectual property, privacy issues or project jeopardizing risks. That is, the data of European laboratories must be *as open as possible and as closed as necessary*.

Open Data has two main goals: first, to promote transparent data in order to achieve a more transparent research and improve the quality of the science; second, to make data open to all citizens in order to encourage innovation and improve the return of investment of science. To make this possible, it is necessary to enhance reusability of data holdings, that is, to plan the dataset management in the long term. It must be noted that good data management is not a goal in itself but a tool to make data useful for researchers and communities. FAIR guidelines are designed to promote the findability, accessibility, interoperability and reusability of datasets.

FAIR principles were reported in 2016 by stakeholders from different backgrounds such as academia, industry, funding agencies or publishers [237]. FAIR guidelines acknowledge that humans and machines face distinct barriers and therefore it puts specific emphasis on machine readability and enhancing ability to automatically find and use data by machines. The FAIR principles can be summarised as follows:

- **Findable:** Machine-readable rich metadata are indispensable for automatic discovery of datasets. That way data and metadata will be easier to find for humans and machines. Data and metadata need a globally unique and persistent identifier. In addition, metadata and data have to be indexable.
- **Accessible:** How and who can access that data must be clearly defined. The protocols to access the data have to be open, free and universally implementable.
- **Interoperable:** Datasets usually will be integrated with other data. Therefore, datasets must be syntactically parseable and semantically understandable.
- **Reusable:** Since the ultimate goal of data is to be reused by other researchers and communities then it must be sufficiently described so it can be reused, replicated or integrated with other data. Moreover, it must be shared with clear and least restrictive licenses as possible and meet the standards of the field.

FAIR data is not always open. Open data must be available to everyone to access, use and share without licenses, copyright or patents. However, FAIR data must be accessible by appropriate people, at an appropriate time, in a appropriate way. For example, when new

experimental data is measured, first it will be accessible by the researcher generator of the data. Then, the researcher shares the raw data with partners from the consortia and that data it will be accessible to them. Finally, data is shared with the public upon publication. During all the process the data was accessible since it was well defined who and how data can be access. Nevertheless, only in the last step the data was open to the public. Moreover, FAIR data has also to be findable, interoperable and reusable. That is, FAIR data does not only focus on the accessibility of the data but also on other key aspects.

The BitMap network decided that the dataset will be open and will follow the standards of FAIR guidelines. From the network, we believe that the publication of this dataset will give a broad picture of the current situation of diffuse optics instruments in Europe and it will help to push the standardization of the field and improve the competitiveness of European Union in biophotonics market.

In the next section, I will speak about standardization and how it can benefit the industrialization and commercialization of biophotonics devices.

## 6.2. Standardization

To translate a lab prototype into its commercialization for clinical use is a path full of pitfalls. There are usually two *valleys of death* in that journey: (1) to get authorization to validate the system in the clinics and (2) to convince clinicians that the systems is reliable enough to use on a regular basis on the clinics. To overcome the first valley a lot of paperwork must be done and security requirements must be fulfilled. Regarding the second valley, it has been one of the principal impediments to adopt optical imaging technologies in the clinics since there have been many difficulties to provide absolute physiological parameters. Even if repeatability of the system is very accurate, the recovered parameters are system specific. Therefore, the adoption of shared procedures for performance assessment of diffuse optics devices, by using international consensus standards, is a must. That is, standardized protocols and phantoms must be used to prove the reliability of biophotonics systems. Once systems have been proven to be reliable in the lab, it can be brought to the clinics. That way research quality will improve and commercialization cost will decrease.

To overcome the previously discussed issue many researchers from the European Union are making a big effort to push standarization of biophotonics devices. For example, during October 2018, an European workshop on *Performance Assessment and Standardization in BioPhotonics* was held in Brussels [167]. Most of the discussion hinged on how to translate biophotonics technologies to the medical industry and when/how standardization can be useful for that goal. Some the conclusions that were reached can be summarised as follows [167]:

- Biophotonics technologies are bringing outstanding possibilities due to chemical specificity and functional information. Moreover, since most of them are non-invasive and potentially scalable it makes them a key technology for the healthcare future.
- The adoption of shared and agreed standards for performance assessment of devices can stimulate the development of novel biophotonic techniques for clinical use by increasing the reliability and reproducibility of results.
- Many researchers from different fields of biophotonics have a strong desire for stan-

standardization as research moves into clinical trials. They reported that this step is critical for European competitiveness.

- To facilitate the use of Open Data can enable the use of Machine Learning technologies due to the supply of validated data. Moreover, it can encourage interdisciplinary research.

### 6.3. The diffuse optics systems campaign

As was explained before, the idea of the campaign is twofold: first, to draw a picture of the current state of European instruments in diffuse optics, that is, to show the different kinds of systems that are being used and to check perform variability in standardized phantoms. It should be pointed out that the goal is not to compare one instrument against other but to have a whole picture of the current state of the field. The second goal is to create an open dataset so researchers working on data processing can test different techniques to retrieve optical properties.

The campaign was carried out by one researcher from BitMap network. This researcher went to 10 institutions and performed measurements using 29 different instruments by applying three internationally agreed protocols (BIP, MEDPHOT and nEUROpt protocol, see next section). This researcher carried with him the same phantoms to every institution, that is, the measured phantoms were the same for all instruments. This huge work was conducted in approximately two months to guarantee that conditions of the experiment were the same in all the cases.

In the next section, I will describe each of the three protocols that were used in the campaign. After, I will give a brief summary of the instruments that took part in the campaign.

#### 6.3.1. The European diffuse optics standards

As was discussed previously, the use of standardized protocols for instrumentation performance assessment is a key step to bring biophotonics devices to the clinics. During the last fifteen years, three well-agreed international protocols have been developed in diffuse optics field [229, 169, 227]. These three protocols were followed during the BitMap campaign. The objectives of each protocol can be summarised as follows:

1. **BIP protocol** [229]: It focuses on the characteristics of the system at its basic level. Therefore, neither samples nor data analysis methods are used during the protocol. In particular, BIP protocol addresses several parameters of the system source, the responsivity of the detection system, the differential nonlinearity of the timing electronics and the stability, shape and background of the instrumental response function (IRF).
2. **MEDPHOT protocol** [169]: It assess the capabilities of instruments to measure homogeneous optical properties ( $\mu_a$  and  $\mu'_s$ ) in turbid media. In this protocol, the system is considered as a black-box and specifications are not taken into consideration, in other words, MEDPHOT is a high-level protocol since it assesses the instrument as a whole, including also the data analysis tools. The different tests done measure the accuracy, linearity, stability, reproducibility and noise on optical properties recovering.

3. **nEUROPt protocol** [227]: It is also a high-level protocol but it assess the capabilities of instruments to measure absorption inhomogeneities in turbid media. It was designed to simulate brain activation in the cortex and to evaluate the ability of systems to detect, localize, and quantify absorption changes in the brain.

The set of phantoms were sequentially sent to all laboratories. In this campaign, only solid phantoms were used due to their stability, reproducibility and easy-to-use characteristics. For BIP protocol responsivity phantom [229] was used. A set of 32 homogeneous solid phantoms [169] that cover a wide range of scattering and absorption properties were used for MEDPHOT protocol. Finally, a solid switchable phantom [170] was used for nEUROPT protocol.

### 6.3.2. The instruments involved in the campaign

The idea of the campaign is to include a wide spectrum of state-of-the-art instruments from European laboratories. The institutions that took part in the campaign are Politecnico di Milano (POLIMI) in Italy, Physikalisch-Technische Bundesanstalt (PTB) in Germany, University Hospital of Birmingham (UHB) in United Kingdom, University of Birmingham (UoB) in United Kingdom, Nalecz Institute of Biocybernetics and Biomedical Engineering (IBIB) in Poland, University College London (UCL) in United Kingdom, Instituto de Ciencias Fotónicas (ICFO) in Spain, Hospital Universitario Valle de Hebrón (VdH) in Spain, ICube in France and University of Strasbourg (UoS) in France.

The participating instruments can be classified using the following characteristics:

- **Development phase:** Some of the instruments are lab prototypes which have not been taken out of the university. Other instruments have already been tested in the clinics and others are commercial instruments with “Conformité Européenne” label or FDA approval.
- **Technologies:** Continuous-wave (CW), frequency domain (FD) and time-resolved (TR) instruments were included in the campaign.
- **Specifications:** The range of specifications is very broad containing for example multi-wavelength instruments (some of them broadband), multi-detector systems (e.g. OCTOPUS system uses eight detectors to perform tomography) and multi-modality system such as LUCA which uses diffuse optics in combination with ultrasound technology.
- **Purpose:** Most of the instruments are oximeters or spectrometers with several applications into the clinics such as adult/newborn brain imagers, hemodynamics monitors, thyroid node analysers, mammographers and fetus imagers. There is also one spectrometer that is used to measure fruit and vegetable quality.
- **Institutions:** All institutions were part of the European Union, precisely from Italy, Germany, United Kingdom, Poland and Spain.

In Table 6.1, the characteristics of each instrument are briefly summarized.

## 6.4. Open Data action organization

The campaign to measure the instruments can be considered as the first part of the BitMap project on standardization. The second part involves to collect all the raw data and to organize it in a format with enough metadata so it follows the FAIR standards. I was in charge of this second part, see Figure 6.1 for a summary of the project timeline.

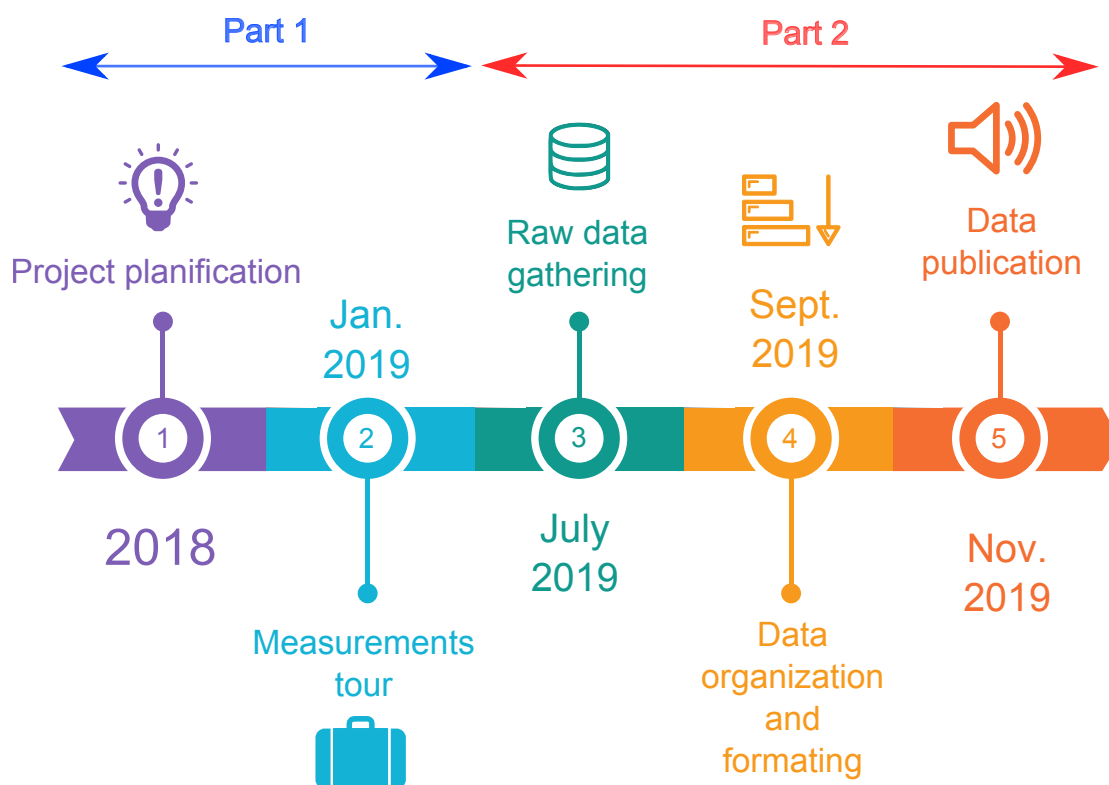


Figure 6.1: Timeline of the whole BitMap standardization and Open Data project.

### 6.4.1. The coordination of institutions

After the first part of the project has been finished, that is, when all the measurements and protocols have been performed, the second part consists on gathering all the data from the different institutions and organizing and publishing it. This is not a trivial aspect since each institution uses different instruments and saves data in different formats. For that reason, what I planned with all partners was to store the data in several MATLAB matrices for each protocol. With this approach, I managed to put an interface between their own data format and SNIRF format (see next section), which made the process easier for them.

The gathering of raw data phase ended at the beginning of July 2019, see Figure 6.1. During September 2019 all the data will be converted to SNIRF format and between October and November the data will be published to the public domain.

### 6.4.2. Shared Near Infrared File (SNIRF) Format Specification

As was discussed previously, the diffuse optics field is experiencing a lack of standardization that is preventing it to fully develop clinical systems. This lack standardization is not only in terms of instruments assessment but also how data is stored (e.g. the byte stream, the header size, etc). Until now, each laboratory was saving raw data in their own format which make difficult and tedious to share it with other colleges. For that reason, several stakeholders from academia and industry both in software and hardware fields have join forces to push a new data sharing format [200].

| ID | System                               | Lab       | Technique | Application  | TRL |
|----|--------------------------------------|-----------|-----------|--------------|-----|
| 1  | Clinical broadband spectrometer      | POLIMI    | TR        | Spectrometer | 5   |
| 2  | 3mm SiPM                             | POLIMI    | TR        | Oximeter     | 3   |
| 3  | HPM based spectrometer               | PTB       | TR        | Spectrometer | 4   |
| 4  | NIRO 200NX                           | UHB/UoB   | CW        | Oximeter     | 8   |
| 5  | ISS OXIPLEX-TS                       | UHB/UoB   | FD        | Oximeter     | 8   |
| 6  | ISS IMAGENT                          | UHB/UoB   | FD        | Oximeter     | 8   |
| 7  | Multiwavelength system               | IBIB      | TR        | Spectrometer | 5/6 |
| 8  | TD-DCS system                        | IBIB      | TR        | Perfusion    | 4   |
| 9  | SRS based CW system                  | UCL       | CW        | Spectrometer | 6   |
| 10 | TRS-DCS FLOWer                       | ICFO/VdH  | TR        | Oximeter     | 7   |
| 11 | MCP based system                     | PTB       | TR        | Spectrometer | 4   |
| 12 | Fetus clinical instrument, HC1       | ICFO/VdH  | TR        | Oximeter     | 6   |
| 13 | Clinical TD oximeter                 | IBIB      | TR        | Oximeter     | 6   |
| 14 | 8 channel system                     | IBIB      | TR        | Oximeter     | 6   |
| 15 | MAESTROS                             | UCL       | TR        | Spectrometer | 4   |
| 16 | LUCA                                 | POLIMI    | TR        | Spectrometer | 6   |
| 17 | TD oximeter                          | POLIMI    | TR        | Oximeter     | 6   |
| 18 | clinical TD oximeter                 | POLIMI    | TR        | Oximeter     | 6   |
| 19 | wearable TD device                   | POLIMI    | TR        | Spectrometer | 5   |
| 20 | Artinis                              | POLIMI    | CW        | Oximeter     | 8   |
| 21 | Mammograph                           | POLIMI    | TR        | Imaging      | 6   |
| 22 | Multiwavelength "Fruit" spectrometer | POLIMI    | TR        | Spectrometer | 4   |
| 23 | OCTOPUS                              | POLIMI    | TR        | Spectrometer | 4   |
| 24 | Laboratory Spectrometer System       | POLIMI    | TR        | Spectrometer | 6   |
| 25 | Laboratory broadband spectrometer    | POLIMI    | TR        | Spectrometer | 6   |
| 26 | Lab TD-DCS                           | POLIMI    | TR        | Spectrometer | 4   |
| 27 | Mammograph MH                        | POLIMI    | TR        | Imaging      | 6   |
| 28 | SFD Imager                           | ICube/UoS | FD        | Imaging      | 4   |
| 29 | APD based Spectrometer               | ICube/UoS | TR        | Spectrometer | 4   |

Table 6.1: Characteristics of the instruments that were used during the campaign. TRL = Technology Readiness Level of the system under test (e.g. TRL3 = component/subsystem, TRL4 = laboratory prototype, TRL5 = clinical prototype, TRL6 = clinical system already demonstrated in clinics, TRL8 = commercial clinical device).

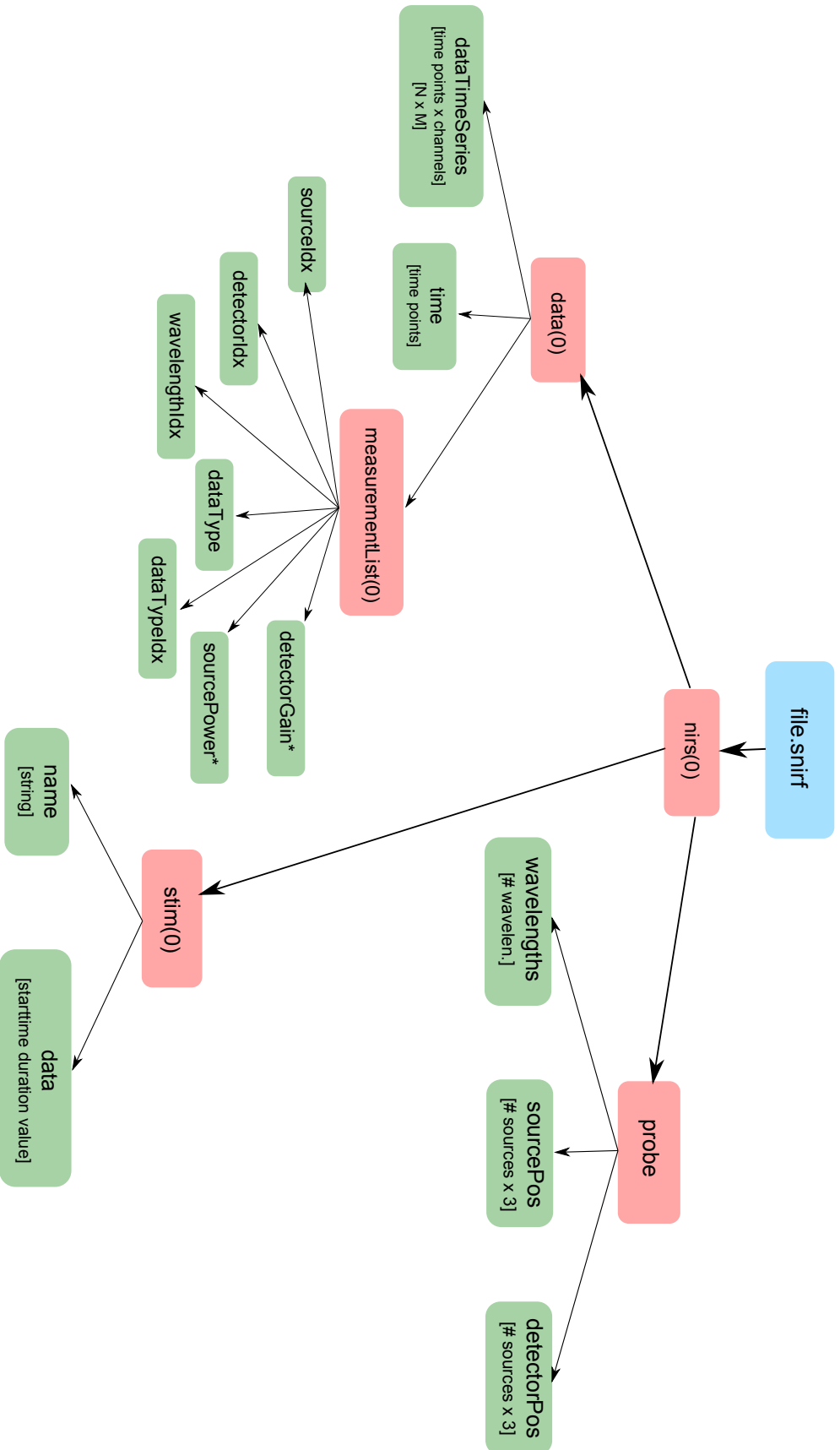


Figure 6.2: Schematic view of SNIRF principal groups (red color) and datasets (green color) adapted for BitMap campaign. The value between the parenthesis indicates the number of the group. In this case, for data there exist only group zero but there could exist as many data groups as necessary. For example, for MEDPHOT protocol there will exist 32 groups since there are 32 different phantoms.

The data format is named Shared Near Infrared File (SNIRF) Format Specification. SNIRF files use the extension `*.snirf` but are HDF5 format files. HDF5 is a format file designed to store, organize and manipulate large amount of data. It can be accessed from different operative systems and there are available many open-source libraries to manipulate HDF5 files from different programming languages, including Java, MATLAB, Python, R and Fortran.

Therefore, HDF5 files are accessible and interoperable in terms of FAIR guidelines. In general terms, HDF5 format uses two primary data organization and storage classes: (1) the group which can be consider as folders and (2) the datasets which are multidimensional arrays of a homogeneous type.

SNIRF format supports measurements from CW, FD and TR technologies including fluorescence measurements. Moreover, Diffuse Correlation Spectroscopy measurements can also be saved in SNIRF format. A schematic view of SNIRF format for BitMap campaign is shown at Figure 6.2.

From `nirs(0)`, three main groups can be distinguished:

- **data group:** It contains one block of measurements. It can extended to many groups by extending the index number. The data contained at the group is:
  - `dataTimeSeries`: Macro-time of the whole experiment.
  - `time`: Micro-time for each DTOF curve.
  - `measurementList`: It contains information about the source, detector and wavelength indexes, the datatype used (e.g. continuous-wave, frequency domain, time-resolved gates and moments) and several characteristics of the instruments. Note that indexes point to the probe group.
- **stim group:** It describes a stimulus condition. For BitMap campaign the stimulus are the phantoms. Several stimulus can be included in one SNIRF file. The data contained in the group is:
  - `name`: String with the name of the phantom.
  - `data`: It defines stimulus time course. In this case, the phantoms are constant over time.
- **probe group:** This group contains the probe geometry. Only one probe group is allowed per SNIRF file. The datasets are:
  - `wavelengths`: The list of wavelengths used.
  - `sourcePos`: Position of the sources.
  - `detectorPos`: Position of the detectors.
  - `frequencies`: List of frequencies used.

It is worth noting that the datasets of the groups described here do not describe the full list. There is a larger set of datasets which are useful for different systems. Here I only described the most common ones.

For the BitMap open dataset one `snirf` file will be created, for each instrument and for each protocol test; for example, in MEDPHOT protocol there will be one `snirf` file for



accuracy/linearity, reproducibility, stability and noise tests. Special mention should be given to how store the IRF of time-resolved instruments. The SNIRF format does not allow to store the IRF of instruments with the measurements. For that reason, I decided to store IRF files in separate `snirf` files. This decision gives enough flexibility since in some tests the IRF was not measured.

### **6.4.3. The open data repository**

The open dataset will be uploaded to Zenodo open-access repository. Zenodo is supported by the European OpenAIRE programme and is managed by CERN [1]. Zenodo encourages to use FAIR principles in the uploaded datasets. For that reason, all the uploaded datasets are automatically assigned a Digital Object Identifier (DOI) to make the content citable, shareable and findable. It is possible to upload newer version of the dataset but original data can never be modified [250].

## **6.5. Conclusions and ongoing work**

In this chapter, I presented the effort that BitMap network have done to push standardization and Open Data in the field of diffuse optics. The first step was to assess the performance of several instruments from European institutions by using three internationally agreed protocols. Those measurements were carried out by the same researcher with the same set of phantoms. The goal of this first step was to draw a global picture of the current state of European labs in diffuse optics and to promote the standardization of instrument performance assessment.

The second part consists on gathering all the data, organize it using the agreed protocol SNIRF and to make it public to the society. The purpose of this second action is to make research more transparent and to encourage the innovation in data analysis. I was in charge of the whole second step and the dataset will be made public at the end of the year.

## Final conclusions and perspectives

The work presented in this thesis had as main goal to study time-resolved diffuse optical tomography algorithms for the estimation of hemoglobin concentration at the human brain cortex.

First, I analyzed the current state-of-the-art of time-resolved diffuse optical tomography theory and algorithms (see Chapter 3). In that chapter, I stressed the drawbacks that commonly used datatypes have in terms of noise correlation and time selectivity. I proposed a novel technique that computes uncorrelated datatypes in the Fourier space. With this new approach, a large set of uncorrelated windows can be used without increasing dramatically the computational time. In this chapter, I proposed to use Tukey or Gaussian windows because of their few oscillations at Fourier space and the decorrelation at time domain. The results on a numerical phantom gave evidence of the superior performance of those windows for inclusions as deep as 3.5 cm due to the improved time selectivity for late times.

In Chapter 4, I tackled another critical aspect of diffuse optical tomography: the inverse problem regularization. In collaboration with the diffuse optics lab from the University of Birmingham, I showed the feasibility of applying total-variation regularization methods into diffuse optical tomography problems defined at irregular meshes. I proposed two different techniques based on finite elements (Section 4.2.1) and graphs approaches (Section 4.2.2). The performed study revealed the improved performance of total-variation techniques compared to Tikhonov regularization. Moreover, within the two proposed methods, graph approach was more stable in coarse meshes than finite elements. This last point, is a critical aspect to take into account when three-dimensional meshes are used since usually fine meshes are not computationally manageable for big models such as an adult human head.

After studying and analyzing different algorithmic aspects of diffuse optical tomography, in Chapter 5, I applied the acquired knowledge into two different in-vivo experiments: arm occlusion (Section 5.3) and motor cortex activation experiments (Section 5.4). The experiments were performed using a time-resolved diffuse optical system, developed at Politecnico di Milano, composed by two probes of three light sources and four silicon photomultiplier detectors each. The system had an optical switcher which switched the optical input among the six light outputs in less than 30 ms. For the experimentation, first arm occlusion experiments were used as validation test for the algorithms; the obtained results were in agreement with previous studies in the literature. After, motor cortex activation experiments were performed

based on finger tapping protocols. Due to the low count rate obtained for two subjects, only results for one subject could be provided. Nevertheless, for this subject, a time-dependant tomographic reconstruction was obtained. Up to my knowledge, the presented time-dependant tomographic reconstructions were never shown in the literature before. These results, are evidence of the feasibility of diffuse optical tomography techniques, to image brain activation with less than one-second time resolution.

The work shown in this thesis brings some new knowledge to the field of diffuse optical tomography. I tried to fill some gaps that I thought were interesting for the community. Nevertheless, as in any scientific work, there are still many questions that need to be answer. In the following paragraphs, I discuss what aspects should be further researched.

Regarding to the theoretical work developed at Chapter 3, I think that there are three points that would be interesting to work on in the future. First, experimental validation have to be performed regarding the extra information that is acquired with the proposed technique. Although, in theory, the information is there, it is still to be proven that the extra information is not lost due to experimental factors. Secondly, it would be interesting to analyze Tukey windows as a limit to full time-resolved tomography reconstructions. In particular, how much new information can be acquired by solving the full DTOF curve. Some question that could arise are: does simulating the full DTOF curve provides new information? Does it make sense to spend computational time on that aspect? Is equivalent information obtained by retrieving the value for some decorrelated datatypes? Lastly, it would be very attractive to build a system that directly measures a gated DTOF. Such system would be considerably faster than state-of-the-art time-resolved systems and, at the same time, the gated DTOF could be used in the proposed reconstruction process since it is equivalent to compute the Tukey windows of a DTOF.

About the performed in-vivo experiments, there are also many aspect to be improved. The system was a lab prototype which was not totally fitted for human brain imaging. It was evident from the count rate values that probe attachment to the scalp was far from optimal in at least two of the subjects. For future studies, a dedicated probe for human brain imaging would ameliorate considerably these aspects and could even provide good count rate values for subjects with dense hair density, see for example [50, 64, 118]. Moreover, the development of a similar time-resolved system with a high density of source and detectors could provide similar results to the gold standard of brain imaging, functional magnetic resonance. Now, with the presented tomographic algorithms, it is feasible to provide clinicians and neuroscientist with time dependant tomographic reconstruction based on diffuse optics technology.

## Publications

### A.1. Published work

- **Wenqi Lu, Jinming Duan, David Orive-Miguel, Lionel Hervé, and Iain B. Styles.** Graph- and finite element-based total variation models for the inverse problem in diffuse optical tomography. *Biomedical Optics Express*, Vol. 10, Issue 6, pp. 2684-2707 (2019). [Journal paper](#).
- **David Orive-Miguel, Lionel Hervé, Jérôme Mars, Laurent Condat, and Pierre Jallon.** Time-resolved diffuse optical tomography: a novel method to compute datatypes allows better absorption quantification. *European Conferences on Biomedical Optics, June 2019, Munich, Germany*. [Talk and proceedings](#).
- **David Orive-Miguel, Pranav Lanka, Lin Yang et al.** The BitMap dataset: an open dataset on performance assessment of diffuse optics instruments. *European Conferences on Biomedical Optics, June 2019, Munich, Germany*. [Talk and proceedings](#).
- **David Orive-Miguel, Lionel Hervé, Jérôme Mars, Laurent Condat, Jean-Marc Dinten, and Sophie Morales.** Computational performance of time-resolved diffuse optical tomography in a two-layer brain model. *Biophotonics congress: biomedical optics congress 2018, April 2018, Miami, USA*. [Poster](#).
- **David Orive-Miguel, Lionel Hervé, Jérôme Mars, Laurent Condat, Jean-Marc Dinten, and Pierre Jallon** Computational performance of Mellin-Laplace transform in neonate brain models. *Optics for neuromonitoring, 2018, Warsaw, Poland*. [Poster](#).
- **Wenqi Lu, David Orive-Miguel, Lionel Hervé, and Iain Styles** Novel graph- and finite element-based total variation models for the inverse problem in diffuse optical tomography. *Optics for neuromonitoring, 2018, Warsaw, Poland*. [Poster](#).
- **Wenqi Lu, Jinming Duan, David Orive-Miguel, Lionel Hervé, and Iain Styles** Novel graph- and finite element-based total variation models for the inverse problem in diffuse optical tomography. *fNIRS conference, 2018, Tokyo, Japan*. [Poster](#).
- **Judy Zouaoui, David Orive-Miguel, Laura Di Sieno, Lionel Hervé, Antonio Pifferi, Andrea Farina, Alberto Dalla Mora, Jacques Derouard, Jerome Mars, Laurent Condat, and Jean-Marc Dinten.** Performance evaluation of time-domain

multispectral diffuse optical tomography in the reflection geometry. *European conferences on biomedical optics, June 2017, Munich, Germany*. [Talk and proceedings](#).

- **David Orive-Miguel, Jean-Marc Dinten, Jerome Mars, Laurent Condat, and Lionel Hervé.** A performance comparison between continuous-wave and time-domain diffuse optical tomography for deep inclusions by using small source-detector separations. *12e Journées imagerie optique non-conventionnelle (JIONC 2017), Mar 2017, Paris, France*. [Talk and proceedings](#).
- **David Orive-Miguel, Lionel Hervé, Jérôme Mars, Laurent Condat, and Jean-Marc Dinten.** Development of time-domain multispectral optical tomography algorithm for brain trauma monitoring. *Journée des doctorants (JDD) EEATS, 2017, Grenoble, France*. [Talk and poster](#).

## A.2. Papers in the pipeline

- **David Orive-Miguel, Lionel Hervé, Laurent Condat, and Jérôme Mars.** A theoretical and numerical analysis of datatypes for time-resolved diffuse optical tomography. [In press. MDPI Applied Sciences journal](#).
- **Collaboration with Politecnico di Milano.** Time-resolved diffuse optical tomography based fNIRS for mapping brain functions. [In writing process](#).
- **Collaboration with BitMap network partners.** The BitMap open dataset on performance assessment of diffuse optics instruments. [In writing process](#).

## Implementation details for photon transport Monte-Carlo method

### B.1. Pseudo-random number generation following Henyey-Greenstein distribution

The inverse probability integral transform is the tool usually used to generate a pseudo-random number that follows distribution  $f(X)$ . As prerequisite it is needed to know its cumulative distribution function  $F_X$ . The idea is to find the operator  $T$  that maps the random variable  $U$  that follows a uniform distribution from zero to one,  $U \sim \text{Unif}[0, 1]$  into the arbitrary distribution  $f(X)$ . To calculate the operator let us take  $u \in \text{Unif}[0, 1]$  so,

$$\begin{aligned}
 F_X(u) &= \Pr(X \leq u) \\
 &= \Pr(T(U) \leq u) \\
 &= \Pr(U \leq T^{-1}(u)) \\
 &= T^{-1}(u) \quad (\text{Because } U \sim \text{Unif}[0, 1]).
 \end{aligned}
 \tag{B.1}$$

Therefore, it follows that,

$$T(u) = F_X^{-1}(u), \tag{B.2}$$

that is, the operator that transform a uniformly distributed random number into an arbitrary distribution is the inverse of the arbitrary distribution cumulative function.

If the change of variable  $w = \cos \theta$  is performed into Equation 2.3 then the new phase function is,

$$p(w) = \frac{1}{2} \frac{1 - g^2}{(1 + g^2 - 2gw)^{3/2}} \tag{B.3}$$

which is normalized,

$$\int_{-1}^1 p(w) dw = 1. \tag{B.4}$$

The cumulative distribution of this phase function is,

$$\begin{aligned} C_p(w) &= \frac{1}{2} \int_{-1}^w \frac{1-g^2}{(1+g^2-2gw)^{3/2}} dw \\ &= \frac{1-g^2}{2g} \left[ (1+g^2-2gw)^{-1/2} - (1+g)^{-1} \right], \end{aligned} \quad (\text{B.5})$$

and its inverse function is,

$$\begin{aligned} v &= C_p(C_p^{-1}(v)) \\ &= \frac{1-g^2}{2g} \left[ (1+g^2-2gC_p^{-1}(v))^{-1/2} - (1+g)^{-1} \right] \\ \Rightarrow C_p^{-1}(v) &= \frac{1}{2g} \left[ 1+g^2 - \left( \frac{1-g^2}{1-g+2gv} \right)^2 \right] = \cos \theta. \end{aligned} \quad (\text{B.6})$$

Therefore, the operator  $C_p^{-1}(v)$  generates a number that follows an Henyey-Greenstein distribution when  $v \in \text{Unif}[0, 1]$ .

## B.2. Photon travelling step size

The probability that a photon travels a distance  $L$  without having an scattering event is given by  $p(L) = \mu_s e^{-\mu_s L}$  (note that the probability density function  $p(L)$  is normalized). The cumulative density function is,

$$C_p(L) = \mu_s \int_0^L e^{-\mu_s x} dx = 1 - e^{-\mu_s L}. \quad (\text{B.7})$$

Using the inverse probability integral transform and being  $v \in \text{Unif}[0, 1]$ ,

$$\begin{aligned} v &= C_p(C_p^{-1}(v)) \\ &= 1 - e^{-\mu_s C_p^{-1}(v)} \\ \Rightarrow C_p^{-1}(v) &= -\frac{\ln(1-v)}{\mu_s} = -\frac{\ln v}{\mu_s}, \end{aligned} \quad (\text{B.8})$$

where the last equality holds since  $v \in \text{Unif}[0, 1]$  and therefore  $(1-v) \in \text{Unif}[0, 1]$ .

# Appendix C

## Brief notes on vector calculus

The divergence theorem states that

$$\int_{\Omega} \nabla \cdot u(\mathbf{x}) \, d\Omega = \int_{\partial\Omega} u(\mathbf{x}) \cdot \vec{n} \, d\partial\Omega \quad (\text{C.1})$$

where  $u(\mathbf{x})$  is a scalar function,  $\mathbf{x} \in \Omega$  and  $\partial\Omega$  denotes the boundary of the domain,  $\Omega$ . The first Green identity can be derived by substituting  $\nabla \cdot u(\mathbf{x})$  by  $\nabla \cdot (u\nabla v)$  where  $v(\mathbf{x})$  is another scalar function, then

$$\int_{\Omega} \nabla \cdot (u\nabla v) \, d\Omega = \int_{\partial\Omega} u \frac{\partial v}{\partial n} \, d\partial\Omega, \quad (\text{C.2})$$

and by applying the identity  $\nabla \cdot (u\nabla v) = \nabla u \cdot \nabla v + u\nabla^2 v$  the *first Green's identity* is retrieved

$$\int_{\Omega} \nabla u \cdot \nabla v + u\nabla^2 v \, d\Omega = \int_{\partial\Omega} u \frac{\partial v}{\partial n} \, d\partial\Omega. \quad (\text{C.3})$$

A new equation is obtained by interchanging  $u$  and  $v$  in Equation (C.3), if that new equation is subtracted from Equation (C.3) then the *Green's second identity* is derived,

$$\int_{\Omega} u\nabla^2 v - v\nabla^2 u \, d\Omega = \int_{\partial\Omega} u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \, d\partial\Omega. \quad (\text{C.4})$$





## Generalized least squares and datatypes

Let us define a random vector  $\mathbf{Z}$  with a covariance matrix  $\text{cov}[\mathbf{Z}]$ . After applying a linear transformation  $\mathbf{A}$  to the random vector the covariance matrix is,

$$\text{cov}[\mathbf{AZ}] = \mathbf{A}\text{cov}[\mathbf{Z}]\mathbf{A}^T, \quad (\text{D.1})$$

where the covariance matrix is always semi-positive definite.

### D.1. From ordinary generalized least squares

When using ordinary least squares it is assumed that the data has noise with equal variance, that is, that the covariance matrix is the identity matrix multiplied by a constant,

$$\mathbf{C} = \sigma^2\mathbf{I}. \quad (\text{D.2})$$

The least squares solution is,

$$\mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{y}. \quad (\text{D.3})$$

However, if the data has uncorrelated noise but with different variance for each measure then weighted least squares (WLS) should be used,

$$\mathbf{x} = (\mathbf{A}^T\mathbf{W}^2\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}^2\mathbf{y}, \quad (\text{D.4})$$

where  $\mathbf{W} = \text{diag}(1/\sigma_i)$  and  $\sigma_i$  the noise standard deviation for acquisition  $i$ -th. Then, WLS makes the covariance matrix identity,  $\mathbf{C} = \mathbf{W}\text{diag}(\sigma_i^2)\mathbf{W}^T = \mathbf{I}$ .

In the worst case data noise can be correlated with other measurements. In this case, the covariance matrix is not diagonal anymore so to make the covariance matrix identity and apply ordinary least squares it must be multiplied by,  $\mathbf{C}^{-1/2}$

$$\mathbf{C}^{-1/2}\mathbf{C}\mathbf{C}^{-1/2} = \mathbf{I}. \quad (\text{D.5})$$

The solution is,

$$\mathbf{x} = (\mathbf{A}^T(\mathbf{C}^{-1/2})^T\mathbf{C}^{-1/2}\mathbf{A})^{-1}\mathbf{A}^T(\mathbf{C}^{-1/2})^T\mathbf{C}^{-1/2}\mathbf{y}. \quad (\text{D.6})$$



# Appendix E

## Diffusion equation and Born approximation revisited

### E.1. Born approximation for diffusion coefficient heterogeneities

A small perturbation is introduced into the diffusion coefficient of Equation 2.33,  $D(\mathbf{x}) \rightarrow D(\mathbf{x}) + \delta D(\mathbf{x})$  where  $\delta D(\mathbf{x}) \ll D(\mathbf{x})$ ; in consequence the solution is going to vary a little bit,  $\phi(\mathbf{x}, t) \rightarrow \phi(\mathbf{x}, t) + \delta\phi(\mathbf{x}, t)$ , where also  $\delta\phi(\mathbf{x}, t) \ll \phi(\mathbf{x}, t)$ .

After introducing the new terms to the diffusion approximation equation

$$\frac{1}{c} \frac{\partial(\phi + \delta\phi)}{\partial t} - \nabla \cdot ((D + \delta D)\nabla(\phi + \delta\phi)) + \mu_a(\phi + \delta\phi) = S(\mathbf{x}, t), \quad (\text{E.1})$$

Simplifying the source term  $S(\vec{r}, t)$  with  $\phi(\mathbf{x}, t)$  terms

$$\frac{1}{c} \frac{\partial\delta\phi(\mathbf{x}, t)}{\partial t} - \nabla \cdot (D\nabla\delta\phi(\mathbf{x}, t)) + \mu_a(\mathbf{x}) \delta\phi(\mathbf{x}, t) = \nabla \cdot (\delta D\nabla(\phi + \delta\phi)). \quad (\text{E.2})$$

The solution to Equation (E.2) is

$$\delta\phi(\mathbf{x}, t) = \int_0^T \int_{\Omega} G^0(\mathbf{x}, \mathbf{x}'; t - t') \nabla \cdot (\delta D\nabla(\phi + \delta\phi)) \, d\mathbf{x}' dt' \quad (\text{E.3})$$

and taking into account that  $\delta\phi \ll \phi$ , the source is a Dirac pulse and measurement is taken at detector position the equation converts to,

$$\begin{aligned} \delta\phi(\mathbf{x}_d, t) &= \int_0^T \int_{\Omega} G^0(\mathbf{x}_d, \mathbf{x}'; t - t') \nabla \cdot (\delta D\nabla G^0(\mathbf{x}', \mathbf{x}_s; t - t')) \, d\mathbf{x}' dt' \\ &= \int_0^T \int_{\Omega} G_d^0 \delta D \nabla^2 G_s^0 \, d\mathbf{x}' dt' + \int_0^T \int_{\Omega} G_d^0 (\nabla\delta D \cdot \nabla G_s^0) \, d\mathbf{x}' dt'. \end{aligned} \quad (\text{E.4})$$

Applying first Green's identity to the first integral,

$$\begin{aligned}
\delta\phi(\mathbf{x}_d, t) &= \int_0^T \int_{\Omega} G_d^0 \delta D \nabla^2 G_s^0 \, d\mathbf{x}' dt' + \int_0^T \int_{\Omega} G_d^0 (\nabla \delta D \cdot \nabla G_s^0) \, d\mathbf{x}' dt' \\
&= \int_0^T \int_{\delta\Omega} G_d^0 \delta D \frac{\partial G_s^0}{\partial n} d\mathbf{x}' - \int_0^T \int_{\Omega} \nabla (G_d^0 \delta D) \cdot \nabla G_s^0 \, d\mathbf{x}' dt' \\
&\quad + \int_0^T \int_{\Omega} G_d^0 (\nabla \delta D \cdot \nabla G_s^0) \, d\mathbf{x}' dt',
\end{aligned} \tag{E.5}$$

where the surface integral is null at far boundaries. Second integral can still be developed,

$$\begin{aligned}
\delta\phi(\mathbf{x}_d, t) &= - \int_0^T \int_{\Omega} \nabla (G_d^0 \delta D) \cdot \nabla G_s^0 \, d\mathbf{x}' dt' + \int_0^T \int_{\Omega} G_d^0 (\nabla \delta D \cdot \nabla G_s^0) \, d\mathbf{x}' dt', \\
&= - \int_0^T \int_{\Omega} \delta D (\nabla G_d^0 \cdot \nabla G_s^0) \, d\mathbf{x}' dt'
\end{aligned} \tag{E.6}$$

which is the Born approximation for diffusion coefficient heterogeneities.

## E.2. Dirac delta shift effect

The diffusion approximation equation with a Dirac delta as a source function is

$$\frac{1}{c} \frac{\partial \phi(\mathbf{x}, \tau)}{\partial \tau} - \nabla \cdot (D \nabla \phi(\mathbf{x}, \tau)) + \mu_a(\mathbf{x}) \phi(\mathbf{x}, \tau) = \delta(\mathbf{x}, \tau), \quad \mathbf{x} \in \Omega, \quad \tau > 0. \tag{E.7}$$

By performing the change of variables  $\tau = t - t'$  where  $t'$  is a constant and by the chain rule,

$$\frac{\partial}{\partial \tau} = \frac{\partial t}{\partial \tau} \frac{\partial}{\partial t} + \frac{\partial t'}{\partial \tau} \frac{\partial}{\partial t'} = \frac{\partial}{\partial t}$$

so the new equation will be

$$\frac{1}{c} \frac{\partial \phi(\mathbf{x}, t - t')}{\partial t} - \nabla \cdot (D \nabla \phi(\mathbf{x}, t - t')) + \mu_a(\mathbf{x}) \phi(\mathbf{x}, t - t') = \delta(\mathbf{x}, t - t'), \quad \mathbf{x} \in \Omega, \quad t - t' > 0. \tag{E.8}$$

# Appendix **F**

## Dispersion calculations for several temporal windows

### F.1. Gaussian window

$$\begin{aligned}
 \int_{-\infty}^{\infty} |g(t)|^2 dt &= \int_{-\infty}^{\infty} \left| e^{-t^2/(2\sigma^2)} \right|^2 dt \\
 &= \int_{-\infty}^{\infty} e^{-t^2/\sigma^2} dt \\
 &= \sigma \int_{-\infty}^{\infty} e^{-x^2} dx \\
 &= \sigma\sqrt{\pi}.
 \end{aligned} \tag{F.1}$$

$$\begin{aligned}
 \int_{-\infty}^{\infty} t^2 |g(t)|^2 dt &= \int_{-\infty}^{\infty} t^2 e^{-t^2/\sigma^2} dt \\
 &= \int_{-\infty}^{\infty} t \left( t \cdot e^{-t^2/\sigma^2} \right) dt \\
 &= \left[ t \cdot \left( -\frac{\sigma^2}{2} e^{-t^2/\sigma^2} \right) \right]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} -\frac{\sigma^2}{2} e^{-t^2/\sigma^2} dt \\
 &= \frac{\sigma^3}{2} \sqrt{\pi}.
 \end{aligned} \tag{F.2}$$

$$\begin{aligned}
 \int_{-\infty}^{\infty} |\mathcal{F}[g(t)]|^2 df &= 2\pi\sigma^2 \int_{-\infty}^{\infty} e^{-4\sigma^2\pi^2 f^2} df \\
 &= 2\pi\sigma^2 \frac{\sqrt{\pi}}{2\sigma\pi} = \sigma\sqrt{\pi}.
 \end{aligned} \tag{F.3}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} f^2 |\mathcal{F}[g(t)]|^2 df &= 2\pi\sigma^2 \int_{-\infty}^{\infty} f^2 e^{-4\sigma^2\pi^2 f^2} df \\
&= 2\pi\sigma^2 \int_{-\infty}^{\infty} f (f \cdot e^{-4\sigma^2\pi^2 f^2}) df \\
&= 2\pi\sigma^2 \left[ f \cdot \left( -\frac{1}{8\sigma^2\pi^2} e^{-4\sigma^2\pi^2 f^2} \right) \right]_{-\infty}^{\infty} - 2\pi\sigma^2 \int_{-\infty}^{\infty} -\frac{1}{8\sigma^2\pi^2} e^{-4\sigma^2\pi^2 f^2} df \\
&= \frac{1}{8\sigma\pi^{3/2}}.
\end{aligned} \tag{F.4}$$

$$D_0(g)D_0(\mathcal{F}[g]) = \frac{\sigma^2}{2} \cdot \frac{1}{8\sigma^2\pi^2} = \frac{1}{16\pi^2}. \tag{F.5}$$

## F.2. Laplacian distribution

$$\begin{aligned}
\int_{-\infty}^{\infty} |g(t)|^2 dt &= \int_{-\infty}^{\infty} |e^{-p|t|}|^2 dt \\
&= 2 \int_0^{\infty} e^{-2pt} dt \\
&= \frac{2}{-2p} [e^{-2pt}]_0^{\infty} \\
&= \frac{1}{p}.
\end{aligned} \tag{F.6}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} t^2 |g(t)|^2 dt &= 2 \int_0^{\infty} t^2 e^{-2pt} dt \\
&= 2 \left( \frac{1}{-2p} [t^2 e^{-2pt}]_0^{\infty} + \frac{2}{2p} \int_0^{\infty} t e^{-2pt} dt \right) \\
&= \frac{2}{p} \int_0^{\infty} t e^{-2pt} dt \\
&= \frac{2}{p} \left( \frac{1}{-2p} [t e^{-2pt}]_0^{\infty} + \frac{1}{2p} \int_0^{\infty} e^{-2pt} dt \right) \\
&= \frac{1}{2p^3}.
\end{aligned} \tag{F.7}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} |\mathcal{F}[g(t)]|^2 df &= \int_{-\infty}^{\infty} \left| \frac{2p}{p^2 + (2\pi f)^2} \right|^2 df \\
&= \frac{4}{(2\pi)^2} \gamma^2 \int_{-\infty}^{\infty} \frac{1}{f^4 + 2\gamma^2 f^2 + \gamma^4} df \quad (\text{where } \gamma = p/(2\pi)) \tag{F.8} \\
&= \frac{4}{(2\pi)^2} \gamma^2 \frac{\pi}{2|\gamma|^3} = \frac{1}{p}.
\end{aligned}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} f^2 |\mathcal{F}[g(t)]|^2 df &= \int_{-\infty}^{\infty} f^2 \left| \frac{2p}{p^2 + (2\pi f)^2} \right|^2 df \\
&= \frac{4}{(2\pi)^2} \gamma^2 \int_{-\infty}^{\infty} \frac{f^2}{f^4 + 2\gamma^2 f^2 + \gamma^4} df \quad (\text{where } \gamma = p/(2\pi)) \\
&= \frac{4}{(2\pi)^2} \gamma^2 \int_{-\infty}^{\infty} \frac{f^2}{(f^2 + \gamma^2)^2} df \\
&= \frac{4}{(2\pi)^2} \gamma^2 \int_{-\infty}^{\infty} \frac{1}{f^2 + \gamma^2} - \frac{\gamma^2}{(f^2 + \gamma^2)^2} df \\
&= \frac{4}{(2\pi)^2} \gamma^2 \left[ \frac{\pi}{\gamma} - \frac{\pi}{2\gamma} \right] \\
&= \frac{p}{(2\pi)^2}.
\end{aligned} \tag{F.9}$$

$$D_0(g)D_0(\mathcal{F}[g]) = \frac{1}{2p^2} \cdot \frac{p^2}{(2\pi)^2} = \frac{1}{8\pi^2}. \tag{F.10}$$

### F.3. Rectangle function

$$\int_{-\infty}^{\infty} |g(t)|^2 dt = \int_{-\infty}^{\infty} |\text{rect}(t/\tau)|^2 dt = \tau. \tag{F.11}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} t^2 |g(t)|^2 dt &= \int_{-\infty}^{\infty} t^2 |\text{rect}(t/\tau)|^2 dt \\
&= 2 \int_0^{\tau/2} t^2 |\text{rect}(t/\tau)|^2 dt \\
&= 2 \int_0^{\tau/2} t^2 dt = 2 \left[ \frac{t^3}{3} \right]_0^{\tau/2} = \frac{\tau^3}{12}.
\end{aligned} \tag{F.12}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} |\mathcal{F}[g(t)]|^2 df &= \int_{-\infty}^{\infty} |\tau \text{sinc}(\pi\tau f)|^2 df \\
&= \tau^2 \int_{-\infty}^{\infty} |\text{sinc}(\pi\tau f)|^2 df \\
&= \frac{\tau}{\pi} \int_{-\infty}^{\infty} |\text{sinc}(x)|^2 dx = \tau.
\end{aligned} \tag{F.13}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} f^2 |\mathcal{F}[g(t)]|^2 df &= \int_{-\infty}^{\infty} f^2 |\tau \text{sinc}(\pi\tau f)|^2 df \\
&= \frac{1}{\pi^3 \tau} \int_{-\infty}^{\infty} x^2 |\text{sinc}(x)|^2 dx \\
&= \text{diverges so it is not square integrable.}
\end{aligned} \tag{F.14}$$



#### F.4. Mellin-Laplace moments

$$\begin{aligned} \int_{-\infty}^{\infty} |g(t)|^2 dt &= \int_{-\infty}^{\infty} |t^n e^{-pt} H(t)|^2 dt \\ &= \int_0^{\infty} t^{2n} e^{-2pt} dt = \frac{(2n)!}{(2p)^{2n+1}}. \end{aligned} \quad (\text{F.15})$$

$$\begin{aligned} \int_{-\infty}^{\infty} t^2 |g(t)|^2 dt &= \int_{-\infty}^{\infty} t^2 |t^n e^{-pt} H(t)|^2 dt \\ &= \int_0^{\infty} t^{2n+2} e^{-2pt} dt = \frac{(2n+2)!}{(2p)^{2n+3}}. \end{aligned} \quad (\text{F.16})$$

$$\begin{aligned} \int_{-\infty}^{\infty} |\mathcal{F}[g(t)]|^2 df &= \int_{-\infty}^{\infty} \left| \frac{n!}{(p + i2\pi f)^{n+1}} \right|^2 df \\ &= \frac{n!^2}{2\pi |p|^{2n+1}} \int_{-\pi/2}^{\pi/2} \frac{\sec^2 x}{|(p + i \tan x)|^{2n+2}} dx \quad (2\pi f = p \tan x) \\ &= \frac{n!^2}{2\pi |p|^{2n+1}} \int_{-\pi/2}^{\pi/2} \cos^{2n} x dx \\ &= \frac{n!^2}{\pi |p|^{2n+1}} \int_0^{\pi/2} \cos^{2n} x dx \\ &= \frac{n!^2}{\pi |p|^{2n+1}} W_{2n}. \end{aligned} \quad (\text{F.17})$$

$$\begin{aligned} \int_{-\infty}^{\infty} f^2 |\mathcal{F}[g(t)]|^2 df &= \int_{-\infty}^{\infty} f^2 \left| \frac{n!}{(p + i2\pi f)^{n+1}} \right|^2 df \\ &= \frac{(n! p)^2}{(2\pi)^3 |p|^{2n+1}} \int_{-\pi/2}^{\pi/2} \frac{\tan^2 x \sec^2 x}{|(p + i \tan x)|^{2n+2}} dx \quad (2\pi f = p \tan x, \quad n \geq 1) \\ &= \frac{(n! p)^2}{(2\pi)^3 |p|^{2n+1}} \int_{-\pi/2}^{\pi/2} \tan^2 x \cos^n x dx \\ &= \frac{(n! p)^2}{(2\pi)^3 |p|^{2n+1}} \int_{-\pi/2}^{\pi/2} \sin^2 x \cos^{2n-2} x dx \\ &= \frac{(n! p)^2}{(2\pi)^3 |p|^{2n+1}} \left( \int_{-\pi/2}^{\pi/2} \cos^{2n-2} x dx - \int_{-\pi/2}^{\pi/2} \cos^{2n} x dx \right) \\ &= \frac{(n! p)^2}{(2\pi)^3 |p|^{2n+1}} (2W_{2n-2} - 2W_{2n}) \quad (n \geq 1). \end{aligned} \quad (\text{F.18})$$

$$D_0(g) = \frac{(2n+2)!}{(2p)^{2n+3}} \cdot \frac{(2p)^{2n+1}}{(2n)!} = \frac{(2n+2)!}{(2n)!} \cdot \frac{1}{(2p)^2}. \quad (\text{F.19})$$

$$\begin{aligned} D_0(\mathcal{F}[g]) &= \frac{(n!p)^2}{(2\pi)^3 |p|^{2n+1}} (2W_{2n-2} - 2W_{2n}) \cdot \frac{\pi |p|^{2n+1}}{n!^2 W_{2n}} \\ &= \frac{p^2}{4\pi^2} \left( \frac{W_{2n-2}}{W_{2n}} - 1 \right) \\ &= \frac{p^2}{4\pi^2} \left( \frac{2n}{2n-1} - 1 \right). \end{aligned} \quad (\text{F.20})$$

$$\boxed{D_0(g)D_0(\mathcal{F}[g]) = \frac{1}{16\pi^2} \frac{(2n+2)!}{(2n)!} \left( \frac{2n}{2n-1} - 1 \right), \quad (n \geq 1).} \quad (\text{F.21})$$

Therefore,

$$\frac{1}{16\pi^2} \frac{(2n+2)!}{(2n)!} \left( \frac{2n}{2n-1} - 1 \right) > \frac{1}{16\pi^2}, \quad (n \geq 1). \quad (\text{F.22})$$



## Fourier transform influence on noise

This approach is inspired by the book written by P.C. Hansen [92]. Let us call  $F \in \mathbb{C}^{M \times M}$  the discrete Fourier transform matrix, which is Hermitian and complex. In the following, we analyze how the noise propagates after applying the Fourier transform.

In the case the data is contaminated by independent Gaussian noise with constant variance, the covariance matrix  $C$  is,

$$C = \sigma^2 I.$$

If the Fourier transform is applied then the covariance matrix transforms to

$$C' = \sigma^2 F F^* = m \sigma^2 I.$$

Now, let us assume that the variance of the noise is identical to the signal value  $s_{exact}$ ; the new covariance matrix is

$$C = \text{diag}(s_{exact}),$$

where  $\text{diag}(u)$  indicates a diagonal matrix whose diagonal contains vector  $u$ . After applying the Fourier transform, the covariance matrix transforms to

$$C' = F \text{diag}(s_{exact}) F^*,$$

and it is known that  $F \text{diag}(s_{exact}) F^*$  is a circulant matrix with its first column as  $F s_{exact}$ , the Fourier transform of  $s_{exact}$ . If the exact signal is dominated by low frequencies, then the largest components of the first column  $F s_{exact}$  will be located at the top and the covariance matrix would be almost diagonal with some nonzero covariance terms near the diagonal. This equation is equivalent to Equation 3.21 in Chapter 3.



# Appendix **H**

## Extra time series results for motor cortex activation experiments

In this appendix the rest of time series for motor cortex activation experiments are included.

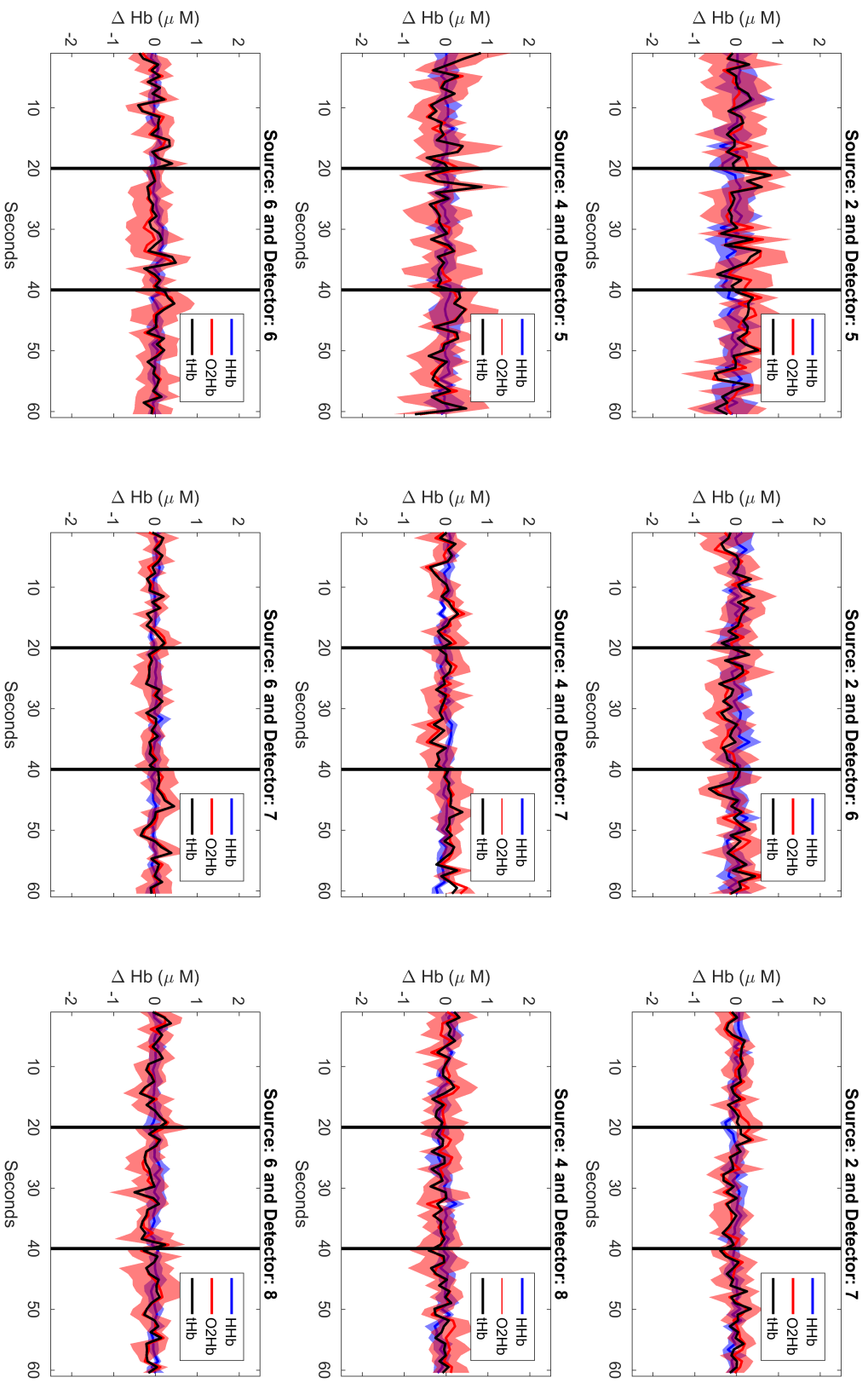


Figure H.1: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for left hemisphere during no hand tapping performed by Subject 1.

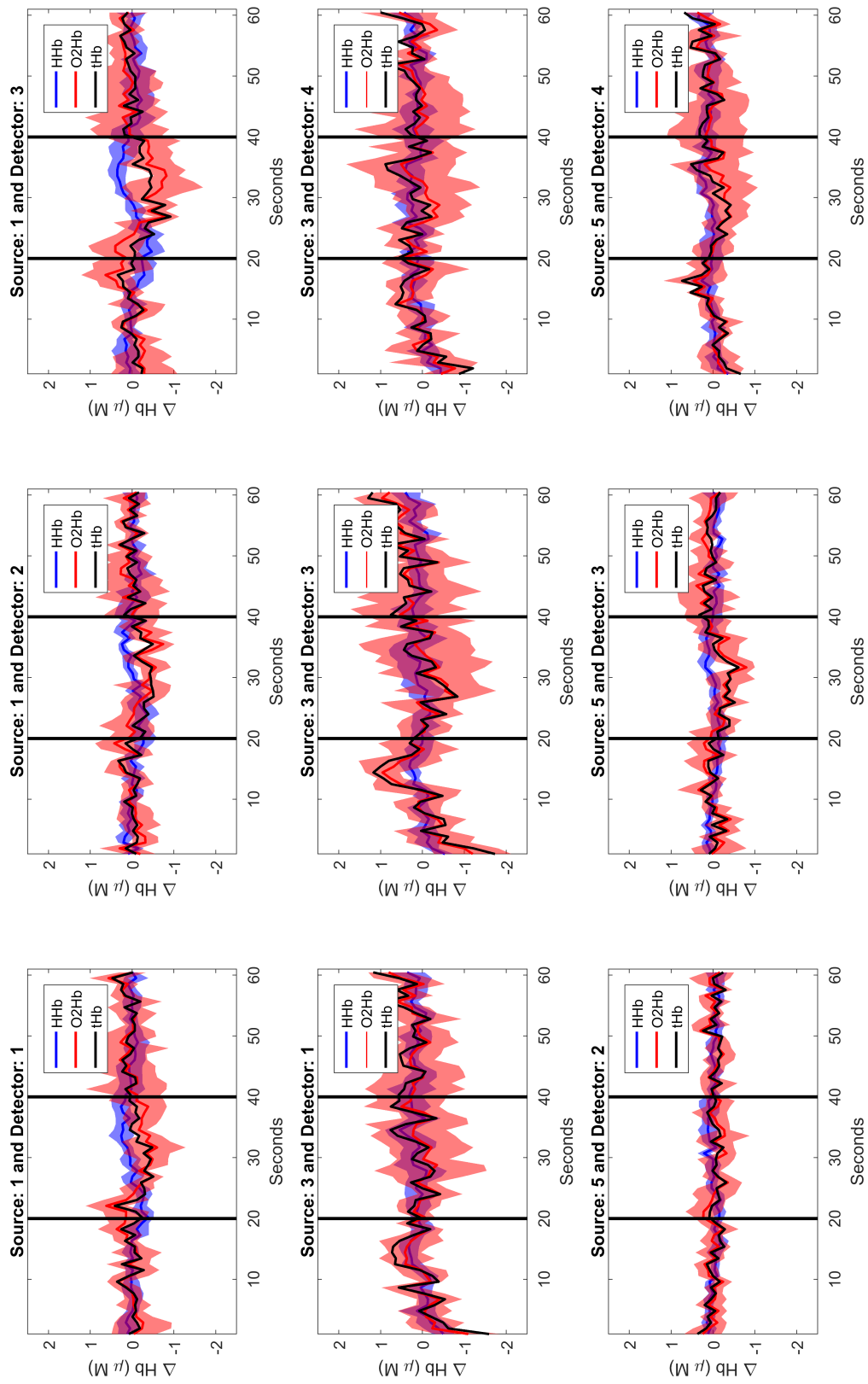


Figure H.2: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for right hemisphere during no hand tapping performed by Subject 1.



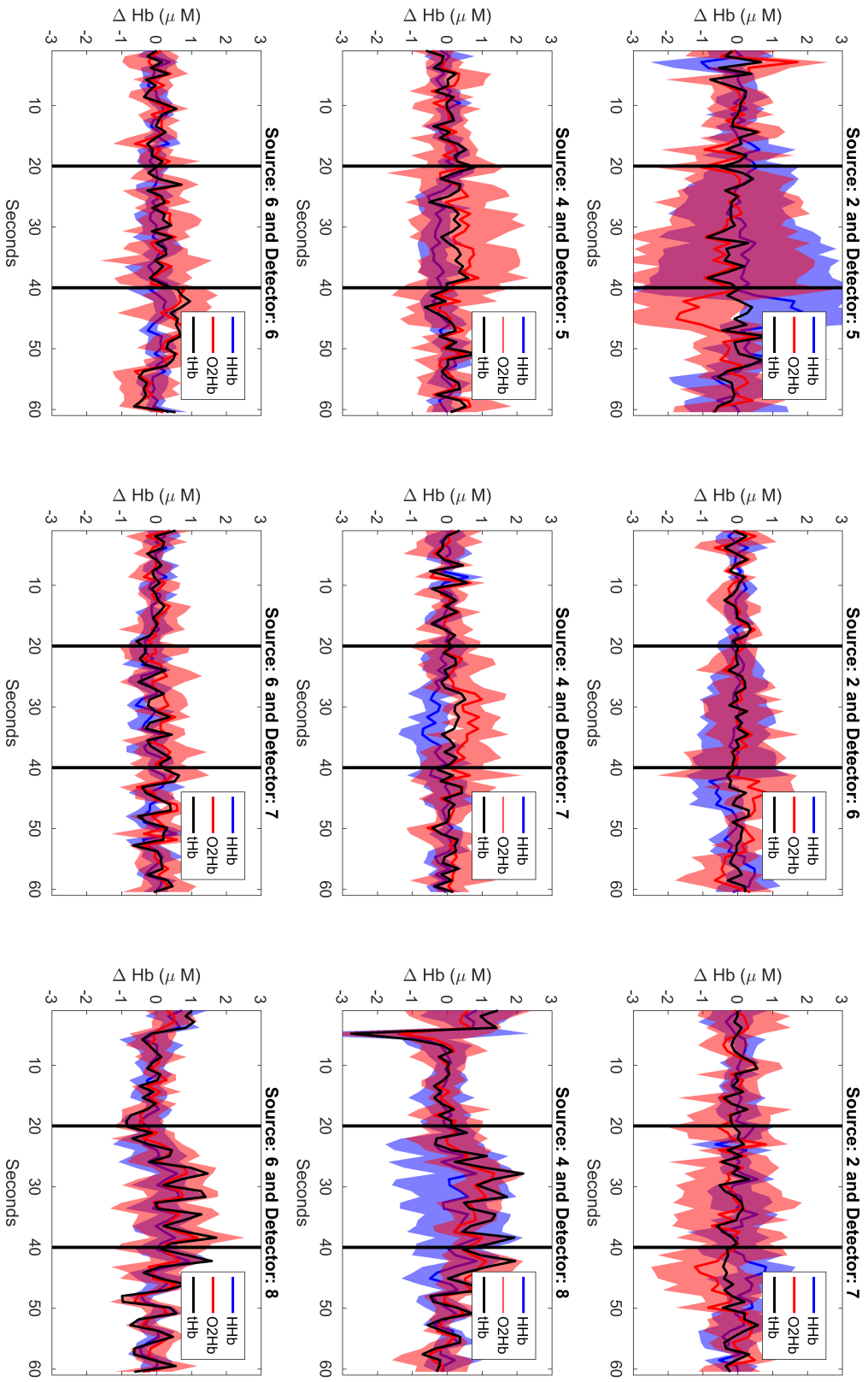


Figure H.3: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for left hemisphere during right hand tapping performed by Subject 2.

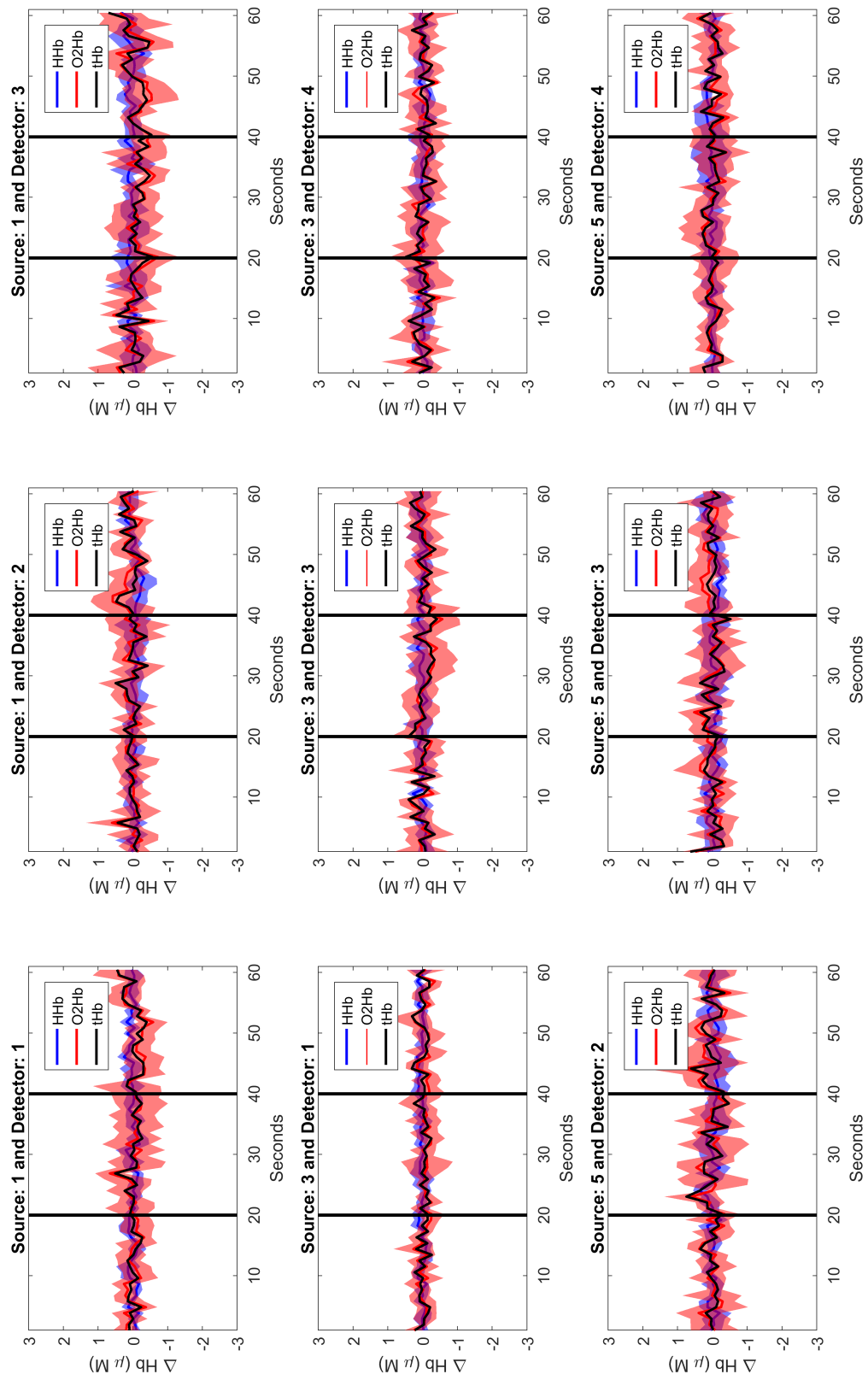


Figure H.4: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for right hemisphere during right hand tapping performed by Subject 2.

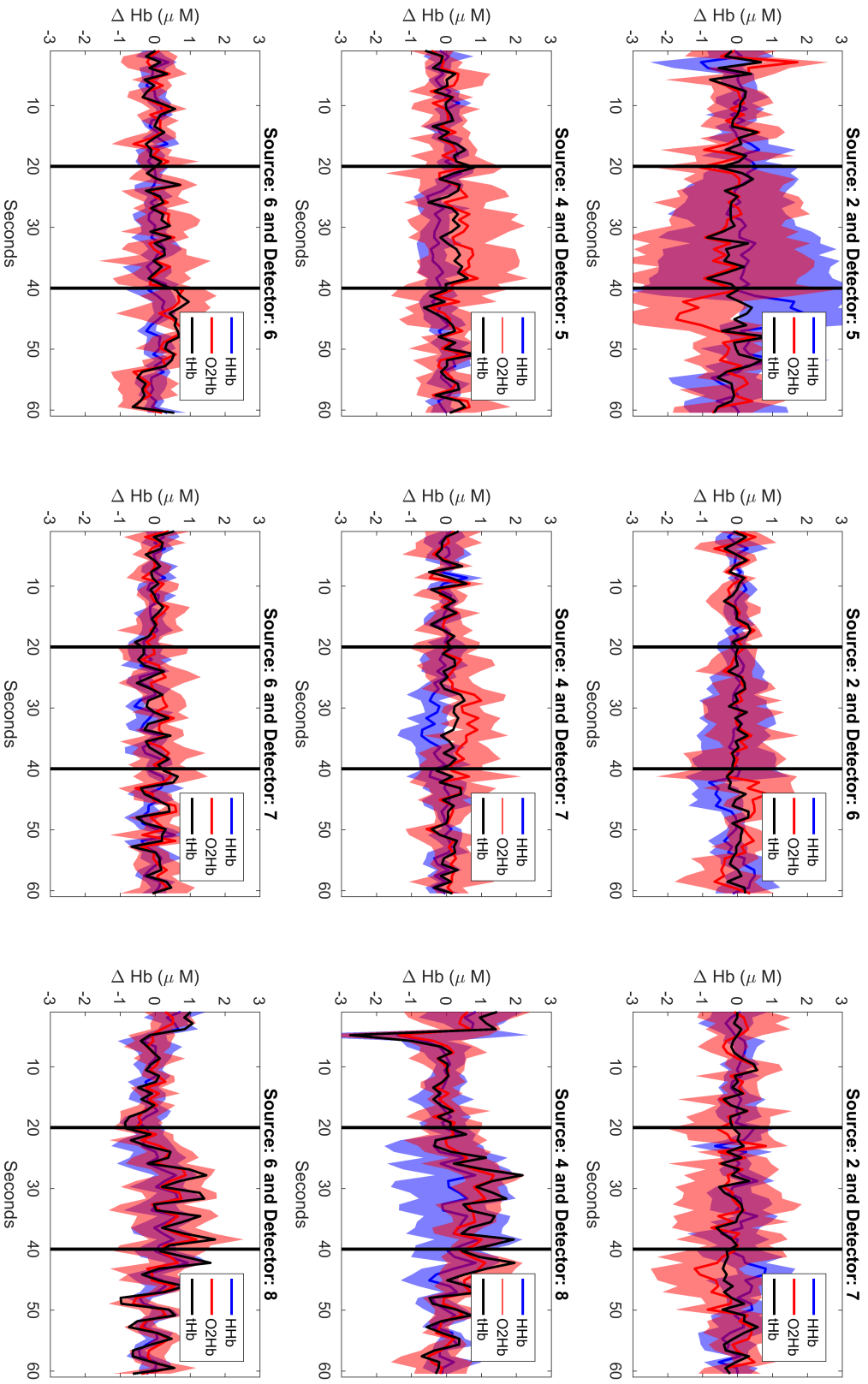


Figure H.5: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for left hemisphere during no hand tapping performed by Subject 2.

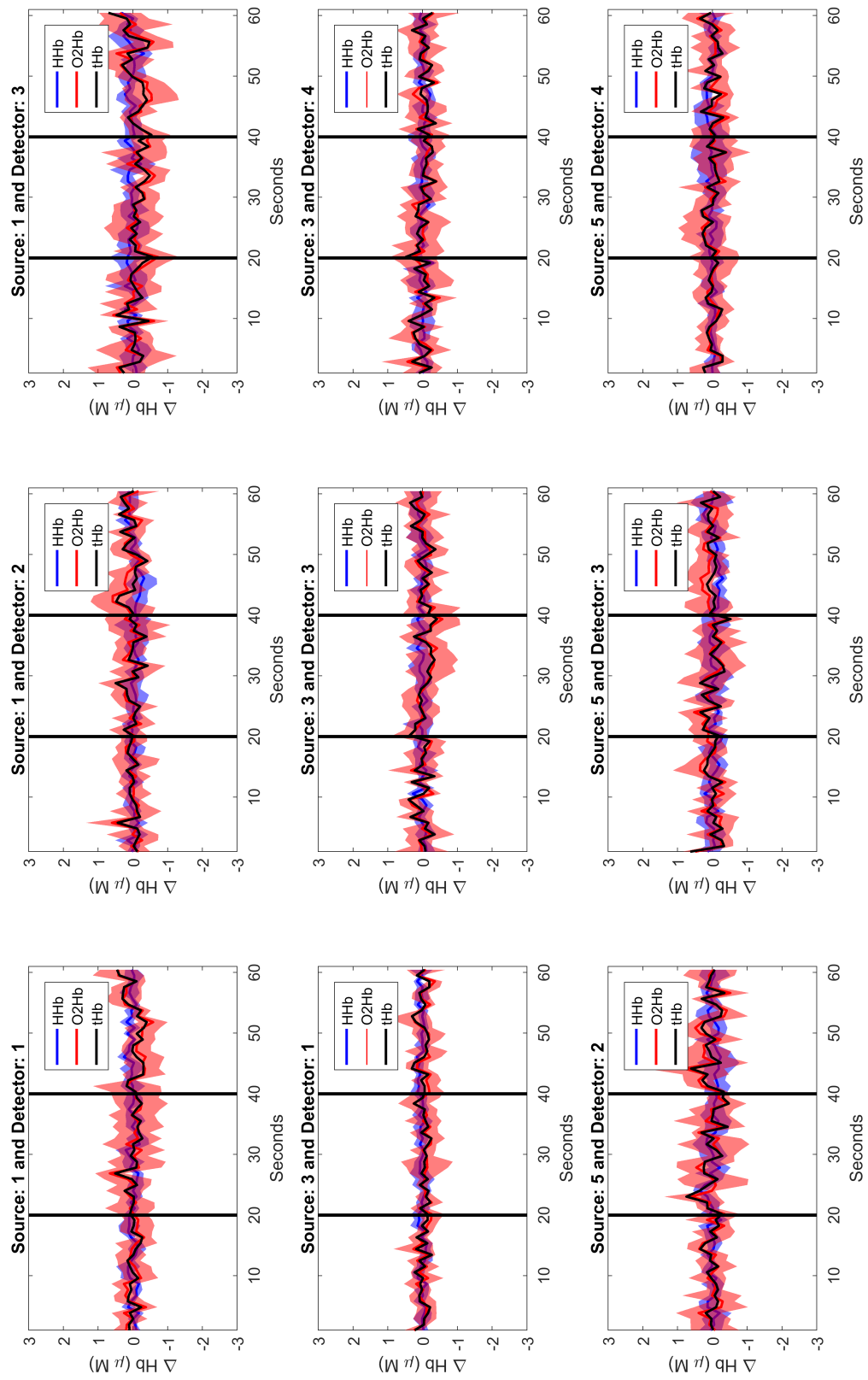


Figure H.6: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for right hemisphere during no hand tapping performed by Subject 2.

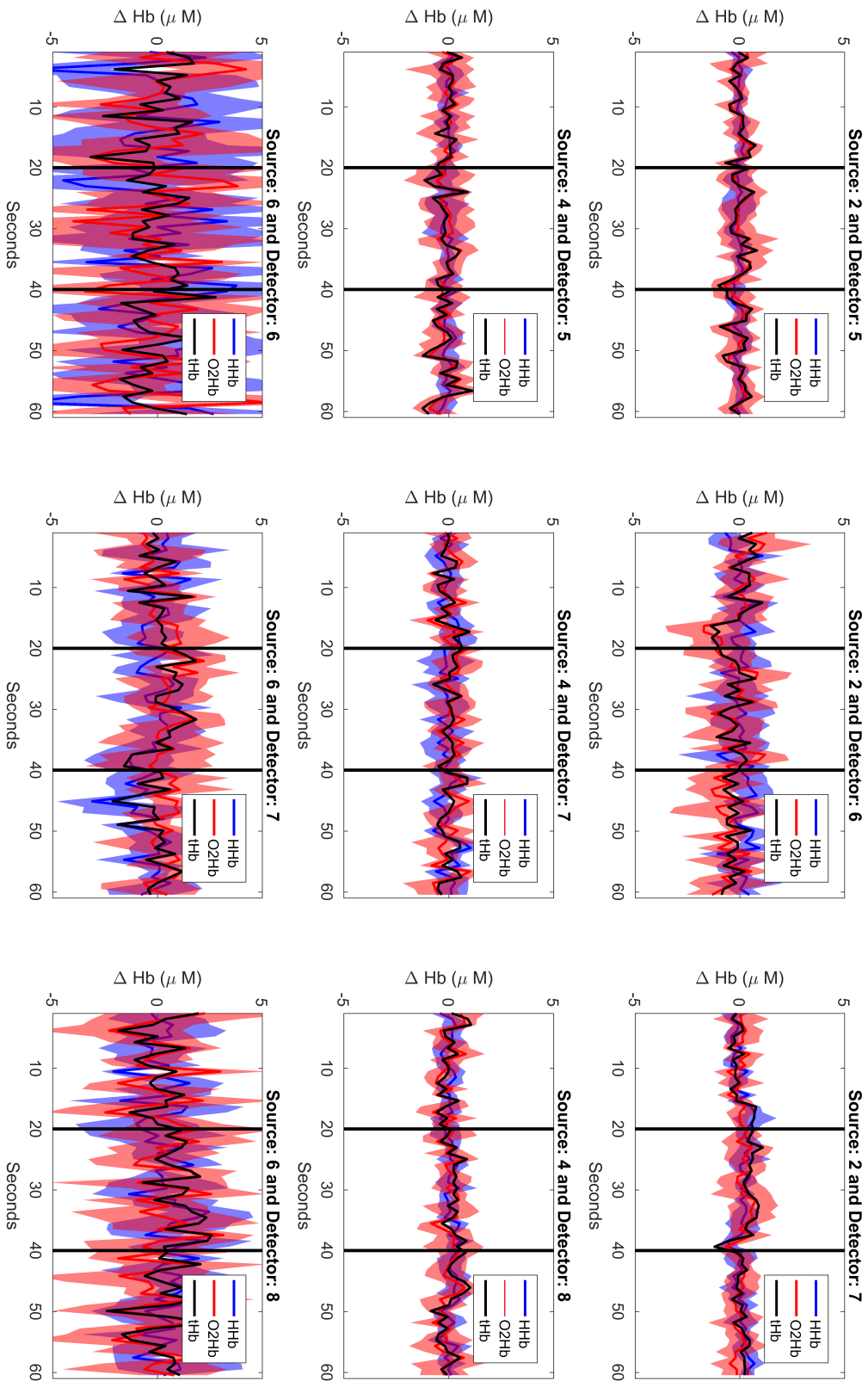


Figure H.7: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for left hemisphere during left hand tapping performed by Subject 3.

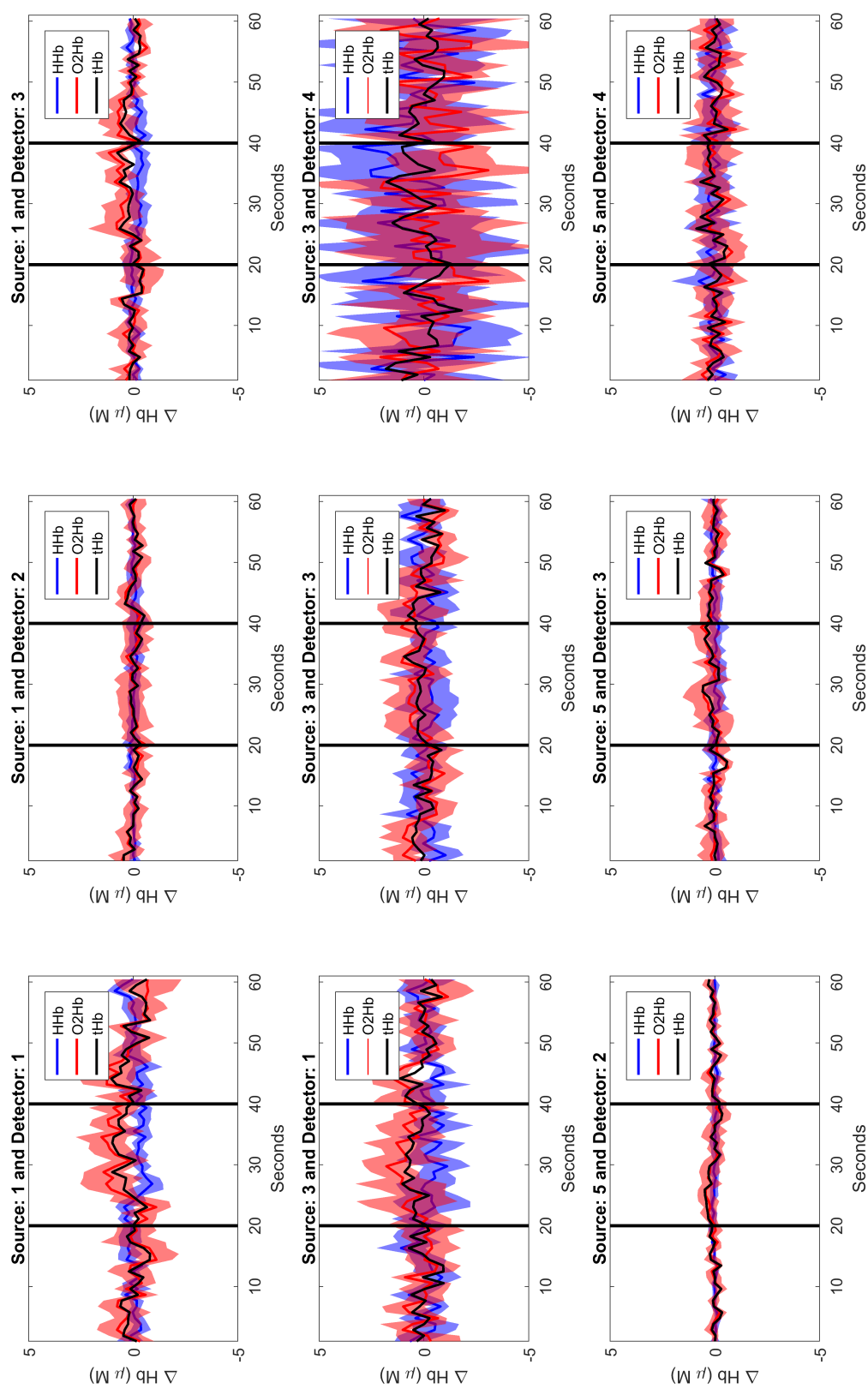


Figure H.8: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for right hemisphere during left hand tapping performed by Subject 3.

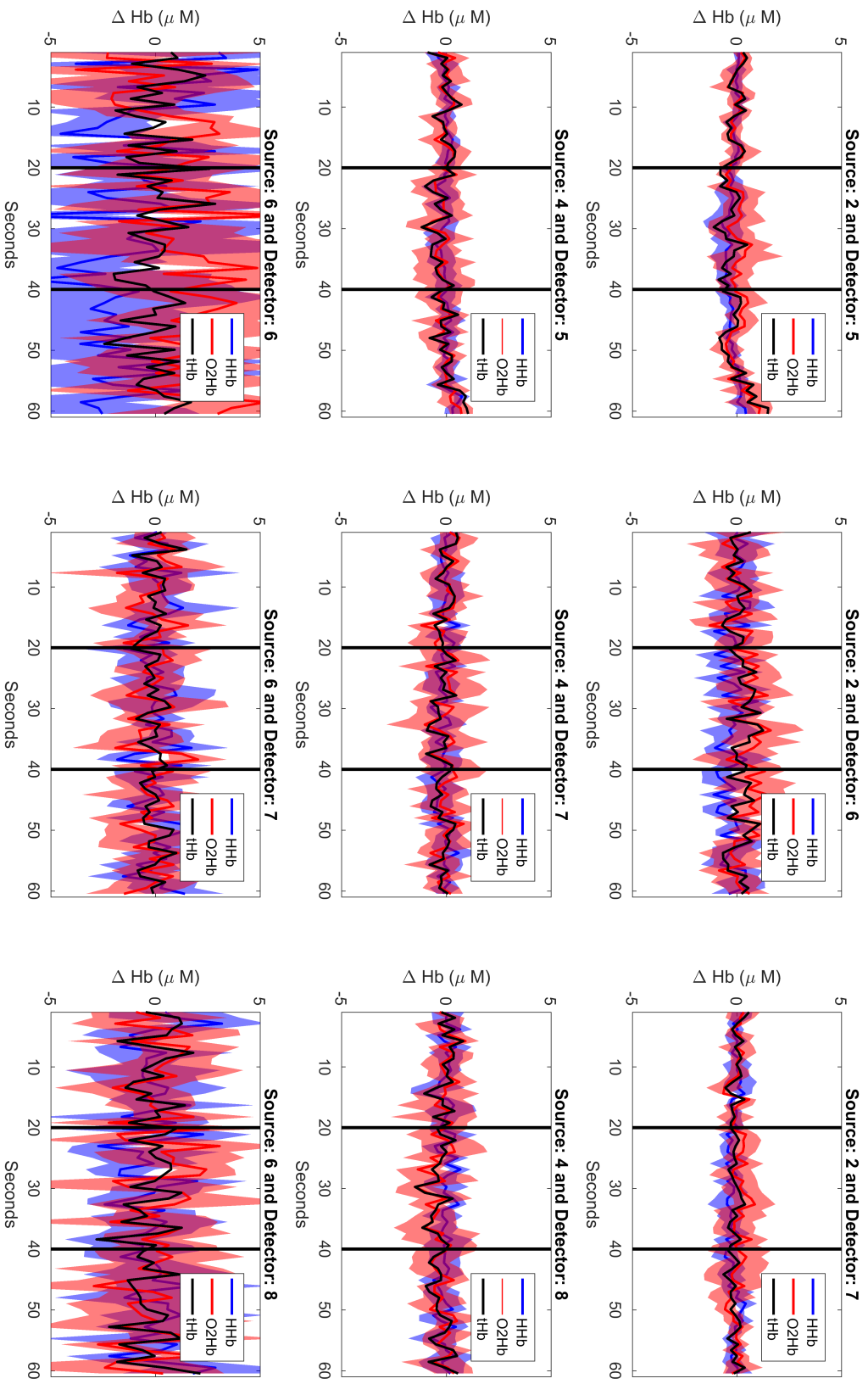


Figure H.9: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for left hemisphere during right hand tapping performed by Subject 3.



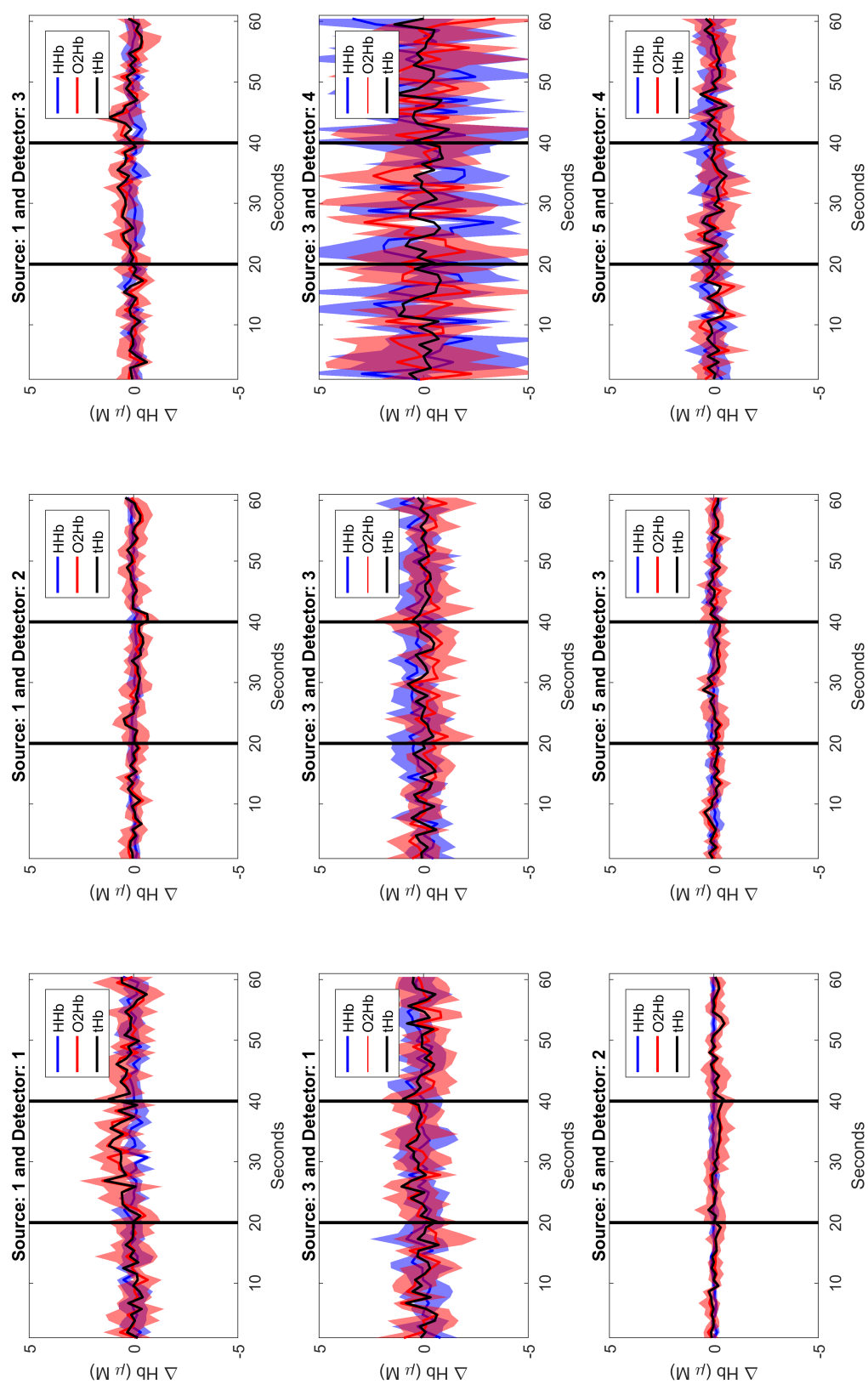


Figure H.10: Time series of changes in oxygenated (O2Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for right hemisphere during right hand tapping performed by Subject 3.



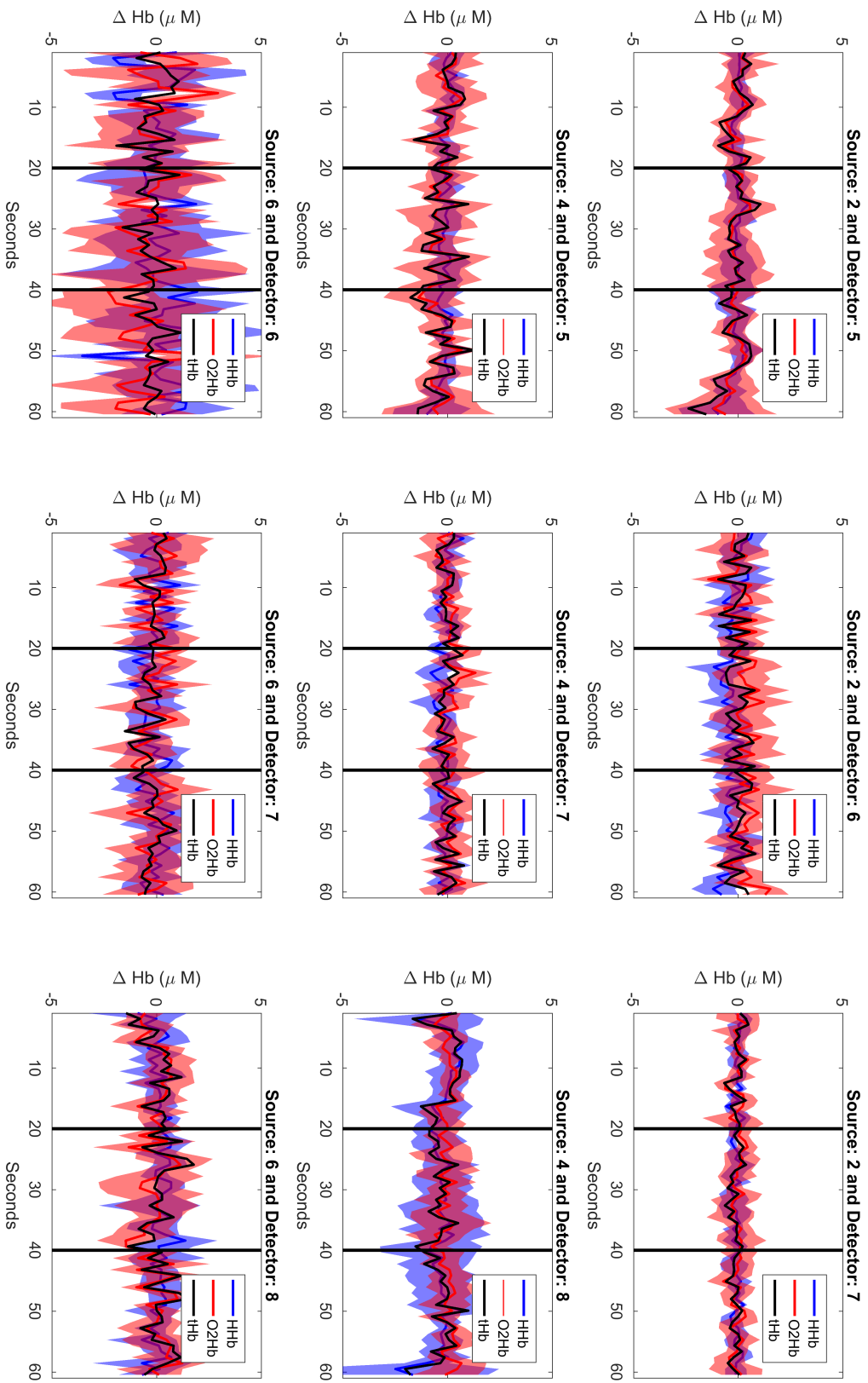


Figure H.11: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for left hemisphere during no hand tapping performed by Subject 3.

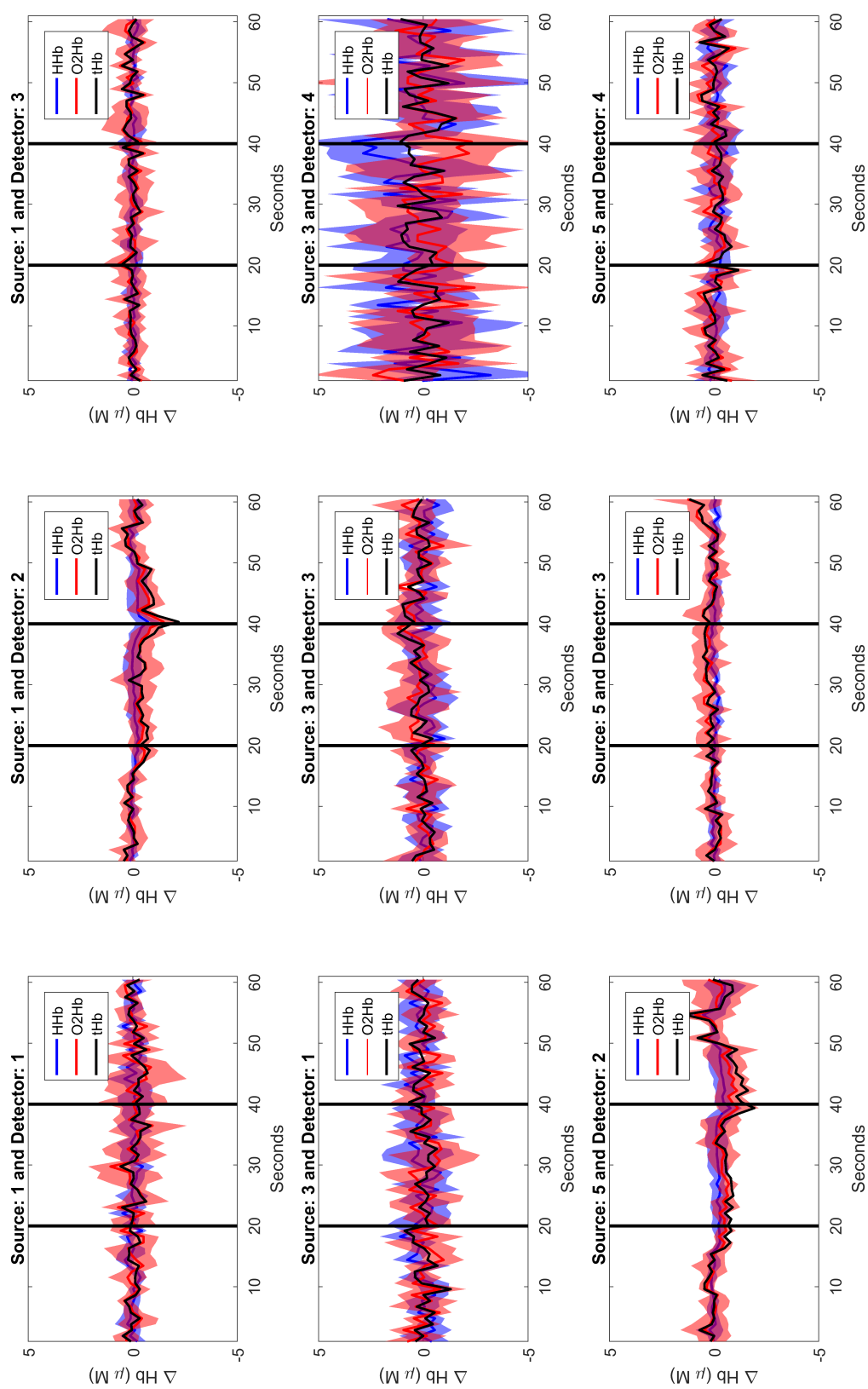


Figure H.12: Time series of changes in oxygenated (O<sub>2</sub>Hb, red line) and deoxygenated (HHb, blue line) hemoglobin concentration for right hemisphere during no hand tapping performed by Subject 3.



# Bibliography

- [1] About Zenodo. <http://about.zenodo.org/>. Accessed: 2019-07-09.
- [2] R. G. Abramson, K. R. Burton, J.-P. J. Yu, E. M. Scalzetti, T. E. Yankeelov, A. B. Rosenkrantz, M. Mendiratta-Lala, B. J. Bartholmai, D. Ganeshan, L. Lenchik, and R. M. Subramaniam. Methods and challenges in quantitative imaging biomarker development. *Academic Radiology*, 22(1):25–32, 2015.
- [3] S. H. Ahn and E. Boros. Nuclear and optical bimodal imaging probes using sequential assembly: A perspective. *Cancer Biotherapy and Radiopharmaceuticals*, 33(8):308–315, 2018.
- [4] E. Alerstam, W. C. Y. Lo, T. D. Han, J. Rose, S. Andersson-Engels, and L. Lilge. Next-generation acceleration and code optimization for light transport in turbid media using GPUs. *Biomed. Opt. Express*, 1(2):658–675, 2010.
- [5] S. R. Arridge. Optical tomography in medical imaging. *Inverse Problems*, 15(2):R41, 1999.
- [6] S. R. Arridge and W. R. B. Lionheart. Nonuniqueness in diffusion-based optical tomography. *Opt. Lett.*, 23(11):882–884, 1998.
- [7] S. R. Arridge and M. Schweiger. Direct calculation of the moments of the distribution of photon time of flight in tissue with a finite-element method. *Appl. Opt.*, 34(15):2683–2687, 1995.
- [8] S. R. Arridge and M. Schweiger. Photon-measurement density functions. Part 2: Finite-element-method calculations. *Appl. Optics*, 34(34):8026–8037, 1995.
- [9] S. R. Arridge, M. Schweiger, and D. T. Delpy. Iterative reconstruction of near-infrared absorption images. In *Inverse Problems in Scattering and Imaging*, volume 1767. SPIE, 1992.
- [10] S. R. Arridge, M. Schweiger, M. Hiraoka, and D. T. Delpy. A finite element approach for modeling photon transport in tissue. *Medical Physics*, 20(2):299–309, 1993.
- [11] O. J. Arthurs, A. Edwards, T. Austin, M. J. Graves, and D. J. Lomas. The challenges of neonatal magnetic resonance imaging. *Pediatric Radiology*, 42(10):1183–1194, 2012.

- [12] T. S. Atabaev, Z. Piao, and A. Molkenova. Carbon dots doped with dysprosium: A bimodal nanoprobe for MRI and fluorescence imaging. *Journal of Functional Biomaterials*, 9(2), 2018.
- [13] G. Bale, S. Mitra, J. Meek, N. Robertson, and I. Tachtsidis. A new broadband near-infrared spectroscopy system for in-vivo measurements of cerebral cytochrome-c-oxidase changes in neonatal brain injury. *Biomed. Opt. Express*, 5(10):3450–3466, 2014.
- [14] B. Baltagi. *Econometrics*. Springer, 2008.
- [15] J. Baritau, K. Hassler, M. Bucher, S. Sanyal, and M. Unser. Sparsity-driven reconstruction for FDOT with anatomical priors. *IEEE Transactions on Medical Imaging*, 30(5):1143–1153, 2011.
- [16] K. Bartley, C. Metayer, S. Selvin, J. Ducore, and P. Buffler. Diagnostic X-rays and risk of childhood leukaemia. *International Journal of Epidemiology*, 39(6):1628–1637, 2010.
- [17] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009.
- [18] A. L. Bertozzi and A. Flenner. Diffuse interface models on graphs for classification of high dimensional data. *Multiscale Modeling & Simulation*, 10(3):1090–1118, 2012.
- [19] Y. Bérubé-Lauzière, M. Crotti, S. Boucher, S. Ettehadi, J. Pichette, and I. Rech. Prospects on time-domain diffuse optical tomography based on time-correlated single photon counting for small animal imaging. *Journal of Spectroscopy*, 2016(1947613):22, 2016.
- [20] P. Bevan. Near-infrared spectroscopy in adult cardiac surgery: An ongoing debate. *Heart, Lung and Circulation*, 24(43), 2015.
- [21] F. Bevilacqua, D. Pigué, P. Marquet, J. D. Gross, B. J. Tromberg, and C. D. Depeursinge. In vivo local determination of tissue optical properties: applications to human brain. *Applied optics*, 38 22:4939–50, 1999.
- [22] I. J. Bigio and S. Fantini. *Quantitative biomedical optics: Theory, methods, and applications*. Cambridge University Press, 1st edition, 2016.
- [23] T. Binzoni, T. S. Leung, A. H. Gandjbakhche, D. Rüfenacht, and D. T. Delpy. The use of the Henyey–Greenstein phase function in Monte Carlo simulations in biomedical optics. *Physics in Medicine & Biology*, 51(17):N313, 2006.
- [24] A. Y. Bluestone, M. Stewart, J. M. Lasker, G. S. Abdoulaev, and A. H. Hielscher. Three-dimensional optical tomographic brain imaging in small animals, part 1: hypercapnia. *Journal of Biomedical Optics*, 9:9 – 9 – 17, 2004.
- [25] A. Y. Bluestone, M. Stewart, B. Lei, I. S. Kass, J. M. Lasker, G. S. Abdoulaev, and A. H. Hielscher. Three-dimensional optical tomographic brain imaging in small animals, part 2: Unilateral carotid occlusion. *Journal of Biomedical Optics*, 9:9 – 9 – 11, 2004.
- [26] J. Boutet, L. Hervé, M. Debourdeau, L. Guyon, P. Peltie, J. Dinten, L. Saroul, F. Duboeuf, and D. Vray. Bimodal ultrasound and fluorescence approach for prostate cancer diagnosis. *Journal of Biomedical Optics*, 14(6):064001, 2009.

- [27] T. Braun, U. Kunz, C. Schulz, A. Lieber, and C. Willy. Near-infrared spectroscopy for the detection of traumatic intracranial hemorrhage: Feasibility study in a German army field hospital in Afghanistan. *Der Unfallchirurg*, 118(8):693–700, 2015.
- [28] X. Bresson, X.-C. Tai, T. F. Chan, and A. Szlam. Multi-class transductive learning based on l1 relaxations of Cheeger cut and Mumford-Shah-Potts model. *Journal of mathematical imaging and vision*, 49(1):191–201, 2014.
- [29] C. Bulgarelli, A. Blasi, S. Arridge, S. Powell, C. de Klerk, V. Southgate, S. Brigadoi, W. Penny, S. Tak, and A. Hamilton. Dynamic causal modelling on infant fNIRS data: A validation study on a simultaneously recorded fNIRS-fMRI dataset. *NeuroImage*, 175:413–424, 2018.
- [30] J. Bunch and L. Kaufman. Some stable methods for calculating inertia and solving symmetric linear systems. *Mathematics of Computation - Math. Comput.*, 31:163–163, 1977.
- [31] M. Buttafava, E. Martinenghi, D. Tamborini, D. Contini, A. D. Mora, M. Renna, A. Torricelli, A. Pifferi, F. Zappa, and A. Tosi. A compact two-wavelength time-domain NIRS system based on SiPM and pulsed diode lasers. *IEEE Photonics Journal*, 9(1):1–14, 2017.
- [32] R. Buxton. *Introduction to functional magnetic resonance imaging: principles and techniques*. Cambridge University Press, 2009.
- [33] R. Buxton, K. Uludağ, D. Dubowitz, and T. Liu. Modeling the hemodynamic response to brain activation. *Neuroimage*, 23:S220–S233, 2004.
- [34] CERN Open Data Portal. <http://opendata.cern.ch/>. Accessed: 2019-01-09.
- [35] M. J. Chan, T. Chung, N. J. Glassford, and R. Bellomo. Near-infrared spectroscopy in adult cardiac surgery patients: A systematic review and meta-analysis. *Journal of Cardiothoracic and Vascular Anesthesia*, 31(4):1155–1165, 2017.
- [36] B. Chance, J. Leigh, H. Miyake, D. Smith, S. Nioka, R. Greenfeld, M. Finander, K. Kaufmann, W. Levy, and M. Young. Comparison of time-resolved and-unresolved measurements of deoxyhemoglobin in brain. *Proceedings of the National Academy of Sciences*, 85(14):4971–4975, 1988.
- [37] Y. Chen, Z. Shen, Z. Shao, P. Yu, and J. Wu. Free flap monitoring using near-infrared spectroscopy: A systemic review. *Annals of Plastic Surgery*, 76(5):590–597, 2016.
- [38] A. Chiarelli, E. Maclin, K. Low, K. Mathewson, M. Fabiani, and G. Gratton. Combining energy and Laplacian regularization to accurately retrieve the depth of brain activity of diffuse optical tomographic data. *Journal of biomedical optics*, 21(3):036008, 2016.
- [39] D. Chitnis, R. J. Cooper, L. Dempsey, S. Powell, S. Quaggia, D. Highton, C. Elwell, J. C. Hebden, and N. L. Everdell. Functional imaging of the human brain using a modular, fibre-less, high-density diffuse optical tomography system. *Biomedical optics express*, 7(10):4275–4288, 2016.
- [40] M. Chu and H. Dehghani. Image reconstruction in diffuse optical tomography based on simplified spherical harmonics approximation. *Opt. Express*, 17(26):24208–24223, 2009.

- [41] V. Coli, E. L. Piccolomini, E. Morotti, and L. Zanni. A fast gradient projection method for 3D image reconstruction from limited tomographic data. *Journal of Physics: Conference Series*, 904(1):012013, 2017.
- [42] L. Condat. Discrete total variation: New definition and minimization. *SIAM Journal on Imaging Sciences*, 10(3):1258–1290, 2017.
- [43] D. Contini, A. Torricelli, A. Pifferi, L. Spinelli, F. Paglia, and R. Cubeddu. Multi-channel time-resolved system for functional near infrared spectroscopy. *Optics Express*, 14(12):5418–5432, 2006.
- [44] R. J. Cooper, E. Magee, N. Everdell, S. Magazov, M. Varela, D. Airantzis, A. P. Gibson, and J. C. Hebden. MONSTIR II: A 32-channel, multispectral, time-resolved optical tomography system for neonatal brain imaging. *Rev. Sci. Instrum.*, 85(5):053105, 2014.
- [45] C. Couprie, L. Grady, L. Najman, J.-C. Pesquet, and H. Talbot. Dual constrained TV-based regularization on graphs. *SIAM Journal on Imaging Sciences*, 6(3):1246–1273, 2013.
- [46] R. Cubeddu, M. Musolino, A. Pifferi, P. Taroni, and G. Valentini. Time-resolved reflectance: a systematic study for application to the optical characterization of tissues. *IEEE Journal of quantum electronics*, 30(10):2421–2430, 1994.
- [47] J. Culver, T. Durduran, D. Furuya, C. Cheung, J. Greenberg, and A. Yodh. Diffuse optical tomography of cerebral blood flow, oxygenation, and metabolism in rat during focal ischemia. *Journal of cerebral blood flow & metabolism*, 23(8):911–924, 2003.
- [48] A. Custo, W. M. Wells, A. Barnett, E. Hillman, and D. A. Boas. Effective scattering coefficient of the cerebral spinal fluid in adult head models for diffuse optical imaging. *Applied optics*, 45 19:4747–55, 2006.
- [49] A. Da Silva, M. Leabad, C. Driol, T. Bordy, M. Debourdeau, J.-M. Dinten, P. Peltié, and P. Rizo. Optical calibration protocol for an x-ray and optical multimodality tomography system dedicated to small-animal examination. *Applied optics*, 48(10):D151–D162, 2009.
- [50] X. Dai, T. Zhang, H. Yang, J. Tang, P. Carney, and H. Jiang. Fast noninvasive functional diffuse optical tomography for brain imaging. *Journal of biophotonics*, 11(3):e201600267, 2018.
- [51] P. Dassonville, X.-H. Zhu, K. Ugurbil, S.-G. Kim, and J. Ashe. Functional activation in motor cortex reflects the direction and the degree of handedness. *Proceedings of the National Academy of Sciences*, 94(25):14015–14018, 1997.
- [52] D. J. Davies, Z. Su, M. T. Clancy, S. Lucas, H. Dehghani, A. Logan, and A. Belli. Near-infrared spectroscopy in the monitoring of adult traumatic brain injury: A review. *Journal of Neurotrauma*, 32:933–941, 2015.
- [53] X. L. Deán-Ben, H. Estrada, and D. Razansky. Shaping volumetric light distribution through turbid media using real-time three-dimensional opto-acoustic feedback. *Opt. Lett.*, 40(4):443–446, 2015.
- [54] H. Dehghani, M. E. Eames, P. K. Yalavarthy, S. C. Davis, S. Srinivasan, C. M. Carpenter, B. W. Pogue, and K. D. Paulsen. Near infrared optical tomography using

- NIRFAST: Algorithm for numerical model and image reconstruction. *Communications in Numerical Methods in Engineering*, 25(6):711–732, 2008.
- [55] H. Dehghani, S. Srinivasan, B. W. Pogue, and A. Gibson. Numerical modelling and image reconstruction in diffuse optical tomography. *Philos. Trans. Royal Soc. A.*, 367(1900):3073–3093, 2009.
- [56] H. Dehghani, B. R. White, B. W. Zeff, A. Tizzard, and J. P. Culver. Depth sensitivity and image reconstruction analysis of dense imaging arrays for mapping brain function with diffuse optical tomography. *Applied optics*, 48(10):D137–D143, 2009.
- [57] D. T. Delpy, M. Cope, P. van der Zee, S. Arridge, S. Wray, and J. Wyatt. Estimation of optical pathlength through tissue from direct time of flight measurement. *Physics in Medicine & Biology*, 33(12):1433, 1988.
- [58] L. Di Sieno, G. Bettega, M. Berger, C. Hamou, M. Aribert, A. Dalla Mora, A. Puszka, H. Grateau, D. Contini, L. Hervé, J.-L. Coll, J.-M. Dinten, A. Pifferi, and A. Planat-Chrétien. Toward noninvasive assessment of flap viability with time-resolved diffuse optical tomography: a preclinical test on rats. *Journal of Biomedical Optics*, 21(2):025004, 2016.
- [59] L. M. Dix, F. van Bel, W. Baerts, and P. M. Lemmers. Comparing near-infrared spectroscopy devices and their sensors for monitoring regional cerebral oxygen saturation in the neonate. *Pediatric Research*, 74:557, 2013.
- [60] R. M. P. Doornbos, R. Lang, M. C. Aalders, F. W. Cross, and H. J. C. M. Sterenborg. The determination of in vivo human tissue optical properties and absolute chromophore concentrations using spatially resolved steady-state diffuse reflectance spectroscopy. *Physics in Medicine & Biology*, 44(4):967, 1999.
- [61] M. Doulgerakis-Kontoudis, A. T. Eggebrecht, S. Wojtkiewicz, J. P. Culver, and H. Dehghani. Toward real-time diffuse optical tomography: accelerating light propagation modeling employing parallel computing on GPU and CPU. *Journal of Biomedical Optics*, 22(12):1 – 11 – 11, 2017.
- [62] D. A. Dunavant. High degree efficient symmetrical Gaussian quadrature rules for the triangle. *International Journal for Numerical Methods in Engineering*, 21(6):1129–1148, 1985.
- [63] T. Durduran, G. Yu, M. G. Burnett, J. A. Detre, J. H. Greenberg, J. Wang, C. Zhou, and A. G. Yodh. Diffuse optical measurement of blood flow, blood oxygenation, and metabolism in a human brain during sensorimotor cortex activation. *Opt. Lett.*, 29(15):1766–1768, 2004.
- [64] A. Eggebrecht, S. Ferradal, A. Robichaux-Viehoever, M. Hassanpour, H. Dehghani, A. Z. Snyder, T. Hershey, and J. Culver. Mapping distributed brain function and networks with diffuse optical tomography. *Nature photonics*, 8(6):448, 2014.
- [65] A. Eggebrecht, D. Muccigrosso, and J. Culver. NeuroDOT: an extensible Matlab toolbox for streamlined optical brain mapping. In *Optical Tomography and Spectroscopy of Tissue XIII*, volume 10874, 2019.
- [66] A. Eklund, P. Dufort, D. Forsberg, and S. M. LaConte. Medical image processing on the GPU – Past, present and future. *Medical Image Analysis*, 17(8):1073 – 1094, 2013.



- [67] A. Elmoataz, O. Lezoray, and S. Boughleux. Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing. *IEEE transactions on Image Processing*, 17(7):1047–1060, 2008.
- [68] Q. Fang. Mesh-based Monte Carlo method using fast ray-tracing in Plücker coordinates. *Biomed. Opt. Express*, 1(1):165–175, 2010.
- [69] Q. Fang and D. A. Boas. Monte Carlo Simulation of Photon Migration in 3D Turbid Media Accelerated by Graphics Processing Units. *Opt. Express*, 17(22):20178–20190, 2009.
- [70] A. Farina, S. Tagliabue, L. Di Sieno, E. Martinenghi, T. Durduran, S. R. Arridge, F. Martelli, A. Torricelli, A. Pifferi, and A. Dalla Mora. Time-domain functional diffuse optical tomography system based on fiber-free silicon photomultipliers. *Applied Sciences*, 7(12):1235, 2017.
- [71] P. Farzam, E. M. Buckley, P.-Y. Lin, K. Hagan, P. E. Grant, T. E. Inder, S. A. Carp, and M. A. Franceschini. Shedding light on the neonatal brain: probing cerebral hemodynamics by diffuse optical spectroscopic methods. *Scientific Reports*, 7(1):15786, 2017.
- [72] E. Ferocino, E. Martinenghi, A. D. Mora, A. Pifferi, R. Cubeddu, and P. Taroni. High throughput detection chain for time domain optical mammography. *Biomed. Opt. Express*, 9(2):755–770, 2018.
- [73] L. Ferreri, E. Bigand, S. Perrey, and A. Bugaiska. The promise of Near-Infrared Spectroscopy (NIRS) for psychological research: a brief review. *L'Annee psychologique*, 114(3):537–569, 2014.
- [74] fNIRS-DOT videos from David Orive-Miguel PhD thesis. <https://zenodo.org/record/3552787>. DOI: 10.5281/zenodo.3552787. Accessed: 2019-11-25.
- [75] M. Forcione, C. Colonnese, and A. Belli. Cerebral hemodynamic influences in task-related functional magnetic resonance imaging and near-infrared spectroscopy in acute sport-related concussion: A review. *Journal of Imaging*, 4(4), 2018.
- [76] M. A. Franceschini, S. Fantini, J. H. Thompson, J. P. Culver, and D. A. Boas. Hemodynamic evoked response of the sensorimotor cortex measured noninvasively with near-infrared optical imaging. *Psychophysiology*, 40(4):548–560, 2003.
- [77] R. D. Frostig. *In vivo optical imaging of brain function*. CRC press, 2009.
- [78] A. Ghosh, C. Elwell, and M. Smith. Cerebral near-infrared spectroscopy in adults: A work in progress. *Anesthesia & Analgesia*, 115(6), 2012.
- [79] P. Giacometti and S. G. Diamond. *Diffuse optical tomography for brain imaging: Continuous wave instrumentation and linear analysis methods*, pages 57–85. Springer New York, New York, NY, 2013.
- [80] A. Gibson, T. Austin, N. Everdell, M. Schweiger, S. R. Arridge, J. Meek, J. Wyatt, D. Delpy, and J. Hebden. Three-dimensional whole-head optical tomography of passive motor evoked responses in the neonate. *Neuroimage*, 30(2):521–528, 2006.
- [81] G. Gilboa and S. Osher. Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, 7:1005–1028, 2008.

- [82] M. S. Gockenbach. *Understanding and implementing the finite element method*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2006.
- [83] G. H. Golub and C. F. Van Loan. *Matrix computations*. Johns Hopkins University Press, Baltimore, MD, USA, 1996.
- [84] P. González-Rodríguez and A. D. Kim. Diffuse optical tomography using the one-way radiative transfer equation. *Biomed. Opt. Express*, 6(6):2006–2021, 2015.
- [85] G. González, V. Kolehmainen, and A. Seppänen. Isotropic and anisotropic total variation regularization in electrical impedance tomography. *Computers & Mathematics with Applications*, 74(3):564 – 576, 2017.
- [86] H. Greenspan, B. van Ginneken, and R. M. Summers. Deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5):1153–1159, 2016.
- [87] C. Groetsch. *Inverse Problems in the Mathematical Sciences*. Springer, 1993.
- [88] J. A. Guggenheim, H. R. A. Basevi, J. Frampton, I. B. Styles, and H. Dehghani. Multi-modal molecular diffuse optical tomography system for small animal imaging. *Measurement Science and Technology*, 24(10):105405, 2013.
- [89] G. Gulsen, B. Xiong, O. Birgul, and O. Nalcioğlu. Design and implementation of a multifrequency near-infrared diffuse optical tomography system. *J. Biomedical Opt.*, 11:11 – 11 – 10, 2006.
- [90] C. Habermehl, S. Holtze, J. Steinbrink, S. P. Koch, H. Obrig, J. Mehnert, and C. H. Schmitz. Somatosensory activation of two fingers can be discriminated with ultrahigh-density diffuse optical tomography. *Neuroimage*, 59(4):3201–3211, 2012.
- [91] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton University Bulletin*, pages 49–52, 1902.
- [92] P. C. Hansen. *Discrete inverse problems: Insight and algorithms*. SIAM, 2010.
- [93] T. Harada, I. Miyai, M. Suzuki, and K. Kubota. Gait capacity affects cortical activation patterns related to speed control in the elderly. *Experimental Brain Research*, 193(3):445–454, 2009.
- [94] F. J. Harris. On the use of windows for harmonic analysis with the discrete Fourier transform. *Proc. IEEE*, 66(1):51–83, 1978.
- [95] A. Hasnain, K. Mehta, X. Zhou, H. Li, and N. Chen. Laplace-domain diffuse optical measurement. *Sci. Rep.*, 8(1):12134, 2018.
- [96] J. Hebden, A. Gibson, T. Austin, R. Yusof, N. Everdell, D. Delpy, S. R. Arridge, J. Meek, and J. Wyatt. Imaging changes in blood volume and oxygenation in the newborn infant brain using three-dimensional optical tomography. *Physics in Medicine & Biology*, 49(7):1117, 2004.
- [97] J. Hebden, A. Gibson, R. Yusof, N. Everdell, E. Hillman, D. Delpy, S. R. Arridge, T. Austin, J. Meek, and J. Wyatt. Three-dimensional optical tomography of the premature infant brain. *Physics in Medicine & Biology*, 47(23):4155, 2002.
- [98] L. G. Henyey and J. L. Greenstein. Diffuse radiation in the galaxy. *Astrophysical Journal*, 93:70–83, 1941.

- [99] M. Herranz and A. Ruibal. Optical imaging in breast cancer diagnosis: The next evolution. *Journal of Oncology*, 2012:863747, 2012.
- [100] L. Hervé, A. Koenig, A. Da Silva, M. Berger, J. Boutet, J. Dinten, P. Peltié, and P. Rizo. Noncontact fluorescence diffuse optical tomography of heterogeneous media. *Applied optics*, 46(22):4896–4906, 2007.
- [101] L. Hervé, A. Puszka, A. Planat-Chrétien, and J.-M. Dinten. Time-domain diffuse optical tomography processing by using the Mellin-Laplace transform. *Appl. Optics*, 51(25):5978–5988, 2012.
- [102] P. Hiltunen. *Regularization methods for diffuse optical technology*. PhD thesis, Aalto University, 2011.
- [103] F. Hosomi, M. Yanagi, Y. Kawakubo, N. Tsujii, S. Ozaki, and O. Shirakawa. Capturing spontaneous activity in the medial prefrontal cortex using near-infrared spectroscopy and its application to schizophrenia. *Scientific reports*, 9(1):5283, 2019.
- [104] G. N. Hounsfield. Computerized transverse axial scanning (tomography): Part 1. Description of system. *The British Journal of Radiology*, 46(552):1016–1022, 1973.
- [105] T. Huppert, J. Barker, B. Schmidt, S. Walls, and A. Ghuman. Comparison of group-level, source localized activity for simultaneous functional near-infrared spectroscopy-magnetoencephalography and simultaneous fNIRS-fMRI during parametric median nerve stimulation. *Neurophotonics*, 4(1):015001, 2017.
- [106] T. J. Huppert, M. S. Allen, S. G. Diamond, and D. A. Boas. Estimating cerebral oxygen metabolism from fMRI with a dynamic multicompartment Windkessel model. *Human brain mapping*, 30(5):1548–1567, 2009.
- [107] C. Ichai, H. Quintard, and J.-C. Orban. *Metabolic Disorders and Critically Ill Patients: From Pathophysiology to Treatment*. Springer, 2017.
- [108] J. Idier. Optimisation par majoration-minimisation pour les problèmes inverses linéaires. Gipsa-Lab seminar slides, 2013.
- [109] X. Intes and B. Chance. Multi-frequency diffuse optical tomography. *Journal of Modern Optics*, 52(15):2139–2159, 2005.
- [110] S. L. Jacques. Optical properties of biological tissues: a review. *Physics in Medicine & Biology*, 58(11):R37, 2013.
- [111] M. Jermyn, H. R. Ghadyani, M. A. Mastanduno, W. D. Turner, S. C. Davis, H. Dehghani, and B. W. Pogue. Fast segmentation and high-quality three-dimensional volume mesh creation from medical images for diffuse optical tomography. *Journal of Biomedical Optics*, 18(8):1 – 11 – 11, 2013.
- [112] F. F. Jöbsis. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*, 198(4323):1264–1267, 1977.
- [113] F. F. Jöbsis. Discovery of the near-infrared window into the body and the early development of near-infrared spectroscopy. *Journal of Biomedical Optics*, 4(4):1264–1267, 1999.

- [114] S. Kahraman, H. Kayali, C. Atabey, F. Acar, and S. Gocmen. The accuracy of near-infrared spectroscopy in detection of subdural and epidural hematomas. *Journal of Trauma-Injury Infection & Critical Care*, 61(6), 2006.
- [115] D. Kane, W. Grassi, R. Sturrock, and P. V. Balint. A brief history of musculoskeletal ultrasound: 'From bats and ships to babies and hips'. *Rheumatology*, 43(7):931–933, 2004.
- [116] V. Kavuri, Z. Lin, F. Tian, and H. Liu. Sparsity enhanced spatial resolution and depth localization in diffuse optical tomography. *Biomedical Optics Express*, 3(5):943–957, 2012.
- [117] S.-G. Kim, J. Ashe, K. Hendrich, J. Ellermann, H. Merkle, K. Ugurbil, and A. Georgopoulos. Functional magnetic resonance imaging of motor cortex: hemispheric asymmetry and handedness. *Science*, 261(5121):615–617, 1993.
- [118] S. Koch, C. Habermehl, J. Mehnert, C. Schmitz, S. Holtze, A. Villringer, J. Steinbrink, and H. Obrig. High-resolution optical functional mapping of the human somatosensory cortex. *Frontiers in neuroenergetics*, 2:12, 2010.
- [119] A. Koeing, L. Hervé, V. Josserand, M. Berger, J. Boutet, A. Da Silva, J. Dinten, P. Peltie, J. Coll, and P. Rizo. In vivo mice lung tumor follow-up with fluorescence diffuse optical tomography. *Journal of Biomedical Optics*, 13(1):011008, 2008.
- [120] A. Koenig, L. Hervé, A. Da Silva, J. Dinten, J. Boutet, M. Berger, I. Texier, P. Peltié, P. Rizo, V. Josserand, et al. Whole body small animal examination with a diffuse optical tomography instrument. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 571(1-2):56–59, 2007.
- [121] V. R. Kondepoti, H. M. Heise, and J. Backhaus. Recent applications of near-infrared spectroscopy in cancer diagnosis and therapy. *Analytical and Bioanalytical Chemistry*, 390(1):125, 2007.
- [122] S. Konugolu Venkata Sekar, A. Farina, A. Dalla Mora, C. Lindner, M. Pagliuzzi, M. Mora, G. Aranda, H. Dehghani, T. Durduran, P. Taroni, and A. Pifferi. Broadband (550–1350 nm) diffuse optical characterization of thyroid chromophores. *Scientific Reports*, 8(1):10015, 2018.
- [123] P. Lanka, L. Yang, D. Orive-Miguel, J. Deepak, S. Tagliabue, A. Sudakou, S. Samaei, M. Forcione, Z. Kovacsova, A. Behera, L. Hervé, T. Durduran, A. Liebert, P. Sawosz, A. Belli, I. Tachtsidis, A. Dalla Mora, H. Dehghani, H. Wabnitz, and A. Pifferi. The BITMAP exercise: a multi-laboratory performance assessment campaign of diffuse optical instrumentation. In *Diffuse Optical Spectroscopy and Imaging VII*, volume 11074, 2019.
- [124] J. Leon-Carrion, J. M. Dominguez-Roldan, U. Leon-Dominguez, and F. Murillo-Cabezas. The Infrascanner, a handheld device for screening in situ for the presence of brain haematomas. *Brain Injury*, 24(10):1193–1201, 2010.
- [125] T. Li, Y. Sun, X. Chen, Y. Zhao, and R. Ren. Noninvasive diagnosis and therapeutic effect evaluation of deep vein thrombosis in clinics by near-infrared spectroscopy. *Journal of Biomedical Optics*, 20:20 – 20 – 3, 2015.

- [126] A. Liebert, H. Wabnitz, D. Grosenick, M. Möller, R. Macdonald, and H. Rinneberg. Evaluation of optical properties of highly scattering media by moments of distributions of times of flight of photons. *Appl. Opt.*, 42(28):5785–5792, 2003.
- [127] A. Liebert, H. Wabnitz, J. Steinbrink, H. Obrig, M. Möller, R. Macdonald, A. Villringer, and H. Rinneberg. Time-resolved multidistance near-infrared spectroscopy of the adult head: intracerebral and extracerebral absorption changes from moments of distribution of times of flight of photons. *Appl. Opt.*, 43(15):3037–3047, 2004.
- [128] A. Liemert and A. Kienle. Analytical solution of the radiative transfer equation for infinite-space fluence. *Phys. Rev. A*, 83:015804, 2011.
- [129] A. Liemert and A. Kienle. Exact and efficient solution of the radiative transport equation for the semi-infinite medium. *Sci. Rep.*, 3(1):2018, 2013.
- [130] A. Liemert, D. Reitzle, and A. Kienle. Analytical solutions of the radiative transport equation for turbid and fluorescent layered media. *Sci. Rep.*, 7(1):3819, 2017.
- [131] E. Lin and A. Alessio. What are the basic concepts of temporal, contrast, and spatial resolution in cardiac CT? *Journal of cardiovascular computed tomography*, 3(6):403–408, 2009.
- [132] C. Lindner, M. Mora, P. Farzam, M. Squarcia, J. Johansson, U. M. Weigel, I. Halperin, F. A. Hanzu, and T. Durduran. Diffuse optical characterization of the healthy human thyroid tissue and two pathological case studies. *PLoS ONE*, 11(1):e0147851, 2016.
- [133] J. Liu, S. Wang, X. Cai, S. Zhou, and B. Liu. Hydrogen peroxide degradable conjugated polymer nanoparticles for fluorescence and photoacoustic bimodal imaging. *Chem. Commun.*, 54:2518–2521, 2018.
- [134] A. Liutkus, D. Martina, S. Popoff, G. Chardon, O. Katz, G. Lerosey, S. Gigan, L. Daudet, and I. Carron. Imaging with nature: Compressive imaging using a multiply scattering medium. *Scientific Reports*, 4:5552, 2014.
- [135] W. Lu, J. Duan, D. Orive-Miguel, L. Herve, and I. B. Styles. Graph- and finite element-based total variation models for the inverse problem in diffuse optical tomography. *Biomed. Opt. Express*, 10(6):2684–2707, 2019.
- [136] W. Ma, W. Zhang, X. Yi, J. Li, L. Wu, X. Wang, L. Zhang, Z. Zhou, H. Zhao, and F. Gao. Time-domain fluorescence-guided diffuse optical tomography based on the third-order simplified harmonics approximation. *Appl. Opt.*, 51(36):8656–8668, 2012.
- [137] BITMAP ITN webpage. <http://www.bitmap-itn.eu/>. Accessed: 2018-09-13.
- [138] F. Martelli, T. Binzoni, A. Pifferi, L. Spinelli, A. Farina, and A. Torricelli. There's plenty of light at the bottom: statistics of photon penetration depth in random media. *Scientific reports*, 6:27057, 2016.
- [139] E. Martinenghi, L. Di Sieno, D. Contini, M. Sanzaro, A. Pifferi, and A. Dalla Mora. Time-resolved single-photon detection module based on silicon photomultiplier: A novel building block for time-correlated measurement systems. *Review of Scientific Instruments*, 87(7):073101, 2016.
- [140] S. Matcher, M. Cope, and D. Delpy. In vivo measurements of the wavelength dependence of tissue-scattering coefficients between 760 and 900 nm measured with time-resolved spectroscopy. *Applied Optics*, 36(1):386(11), 1997.

- [141] V. Mattay, J. Callicott, A. Bertolino, A. Santha, J. Van Horn, K. Tallent, J. Frank, and D. Weinberger. Hemispheric control of motor function: a whole brain echo planar fMRI study. *Psychiatry Research: Neuroimaging*, 83(1):7–22, 1998.
- [142] M. Mazurenka, L. Di Sieno, G. Boso, D. Contini, A. Pifferi, A. Dalla Mora, A. Tosi, H. Wabnitz, and R. Macdonald. Non-contact in vivo diffuse optical imaging using a time-gated scanning system. *Biomedical optics express*, 4(10):2257–2268, 2013.
- [143] D. J. McCabe, A. Tajalli, D. R. Austin, P. Bondareff, I. A. Walmsley, S. Gigan, and B. Chatel. Spatio-temporal focusing of an ultrafast pulse through a multiply scattering medium. *Nature Communications*, 2:447, 2011.
- [144] D. J. Mehagnoul-Schipper, B. van der Kallen, W. N. J. M. Colier, M. C. van der Sluijs, L. van Erning, H. Thijssen, B. Oeseburg, W. Hoefnagels, and R. Jansen. Simultaneous measurements of cerebral oxygenation changes during brain activation by near-infrared spectroscopy and functional magnetic resonance imaging in healthy young and elderly subjects. *Human brain mapping*, 16(1):14–23, 2002.
- [145] E. Merkurjev, T. Kostic, and A. L. Bertozzi. An MBO scheme on graphs for classification and image processing. *SIAM Journal on Imaging Sciences*, 6(4):1903–1930, 2013.
- [146] R. C. Mesquita, T. Durduran, G. Yu, E. M. Buckley, M. N. Kim, C. Zhou, R. Choe, U. Sunar, and A. G. Yodh. Direct measurement of tissue blood flow and metabolism with diffuse optics. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 369(1955):4390–4406, 2011.
- [147] M. Mihara, I. Miyai, M. Hatakenaka, K. Kubota, and S. Sakoda. Sustained pre-frontal activation during ataxic gait: A compensatory mechanism for ataxic stroke? *NeuroImage*, 37(4):1338 – 1345, 2007.
- [148] A. B. Milstein, J. J. Stott, S. Oh, D. A. Boas, R. P. Millane, C. A. Bouman, and K. J. Webb. Fluorescence optical diffusion tomography using multiple-frequency data. *J. Opt. Soc. Am. A*, 21(6):1035–1049, 2004.
- [149] S. Mitra, G. Bale, S. Mathieson, C. Uria-Avellanal, J. Meek, I. Tachtsidis, and N. J. Robertson. Changes in cerebral oxidative metabolism during neonatal seizures following hypoxic–ischemic brain injury. *Frontiers in Pediatrics*, 4:83, 2016.
- [150] B. Montcel, R. Chabrier, and P. Poulet. Detection of cortical activation with time-resolved diffuse optical methods. *Applied optics*, 44(10):1942–1947, 2005.
- [151] Monte-Carlo eXtreme: GPU-based. <http://mcx.space/wiki/>. Accessed: 2019-05-24.
- [152] A. D. Mora, E. Martinenghi, D. Contini, A. Tosi, G. Boso, T. Durduran, S. Arridge, F. Martelli, A. Farina, A. Torricelli, and A. Pifferi. Fast silicon photomultiplier improves signal harvesting and reduces complexity in time-domain diffuse optics. *Opt. Express*, 23(11):13937–13946, 2015.
- [153] M. A. Morris, B. Saboury, B. Burkett, J. Gao, and E. L. Siegel. Reinventing radiology: Big Data and the future of medical imaging. *Journal of Thoracic Imaging*, 33(1):4–16, 2018.

- [154] J. R. Mourant, T. Fuselier, J. Boyer, T. M. Johnson, and I. J. Bigio. Predictions and measurements of scattering and absorption over broadwavelength ranges in tissue phantoms. *Appl. Opt.*, 36(4):949–957, 1997.
- [155] J. Murkin and M. Arango. Near-infrared spectroscopy as an index of brain and tissue oxygenation. *British Journal of Anaesthesia*, 103:i3 – i13, 2009.
- [156] A. Müller, S. Marschall, O. B. Jensen, J. Fricke, H. Wenzel, B. Sumpf, and P. E. Andersen. Diode laser based light sources for biomedical applications. *Laser & Photonics Reviews*, 7(5):605–627, 2012.
- [157] A. Nenna, R. Barbato, S. Matteo Greco, G. Pugliese, M. Lusini, E. Covino, and M. Chello. Near-infrared spectroscopy in adult cardiac surgery: Between conflicting results and unexpected uses. *Journal of Geriatric Cardiology*, 14:659–661, 2017.
- [158] H. Nien and J. A. Fessler. Relaxed linearized algorithms for faster X-ray CT image reconstruction. *IEEE Transactions on Medical Imaging*, 35(4):1090–1098, 2016.
- [159] H. Niu, Z. Lin, F. Tian, S. Dhamne, and H. Liu. Comprehensive investigation of three-dimensional diffuse optical tomography with depth compensation algorithm. *Journal of biomedical optics*, 15(4):046005, 2010.
- [160] H. Niu, F. Tian, Z. Lin, and H. Liu. Development of a compensation algorithm for accurate depth localization in diffuse optical tomography. *Optics letters*, 35(3):429–431, 2010.
- [161] M. Nixon, O. Katz, E. Small, Y. Bromberg, A. A. Friesem, Y. Silberberg, and N. Davidson. Real-time wavefront-shaping through scattering media by all optical feedback. *Nature Photonics*, 7:919–924, 2013.
- [162] S. Okawa, Y. Hoshi, and Y. Yamada. Improvement of image quality of time-domain diffuse optical tomography with lp sparsity regularization. *Biomedical Optics Express*, 2(12):3334–3348, 2011.
- [163] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An Iterative Regularization Method for Total Variation-Based Image Restoration. *Multiscale Modeling & Simulation*, 4(2):460–489, 2005.
- [164] J.-H. Park, W. Sun, and M. Cui. High-resolution in vivo imaging of mouse brain through the intact skull. *Proceedings of the National Academy of Sciences*, 112(30):9236–9241, 2015.
- [165] J.-H. Park, Z. Yu, K. Lee, P. Lai, and Y. Park. Perspective: Wavefront shaping techniques for controlling multiple light scattering in biological tissues: Toward in vivo applications. *APL Photonics*, 3(10):100901, 2018.
- [166] W. Penny, K. Friston, J. Ashburner, S. Kiebel, and T. Nichols. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Elsevier Science, 2011.
- [167] Performance Assessment and Standardization in Biophotonics Workshop. <https://ec.europa.eu/digital-single-market/en/news/performance-assesement-and-standardization-biophotonics-workshop>. Accessed: 2019-07-08.
- [168] Photonics 21. Photonics – a critical Key Enabling Technology for Europe. Role and impact of Photonics in H2020. *Photonics 21*, 2017.

- [169] A. Pifferi, A. Torricelli, A. Bassi, P. Taroni, R. Cubeddu, H. Wabnitz, D. Grosenick, M. Möller, R. Macdonald, J. Swartling, T. Svensson, S. Andersson-Engels, R. L. P. van Veen, H. Sterenborg, J.-M. Tualle, H. Nghiem, S. Avrillier, M. Whelan, and H. Stamm. Performance assessment of photon migration instruments: the MEDPHOT protocol. *Appl. Opt.*, 44(11):2104–2114, 2005.
- [170] A. Pifferi, A. Torricelli, R. Cubeddu, G. Quarto, R. Re, S. Sekar, L. Spinelli, A. Farina, F. Martelli, and H. Wabnitz. Mechanically switchable solid inhomogeneous phantom for performance tests in diffuse imaging and spectroscopy. *Journal of Biomedical Optics*, 20(12):121304, 2015.
- [171] M. A. Pinsky. *Introduction to Fourier analysis and wavelets*, volume 102. Amer. Math. Soc., 2002.
- [172] A. Planat-Chrétien, M. Berger, L. Hervé, L. Watroba, J. Demilly, J. Flament, L. Stimmer, P. Aubourg, and J.-M. Dinten. Time Resolved optical detection for white matter lesion detection: preclinical tests on macaque brains and MRI co-registration. In *European Conference on Biomedical Optics*, page 953806. Optical Society of America, 2015.
- [173] A. Planat-Chrétien, A. Dot, M. Perriollat, M. Berger, R. Lartizien, J.-L. Coll, and G. Bettega. Non-invasive assessment of deep buried flap viability with time-resolved optical monitoring: Results on pigs. Poster, 2018.
- [174] B. Pogue, T. McBride, J. Prewitt, U. Österberg, and K. Paulsen. Spatially variant regularization improves diffuse optical tomography. *Applied optics*, 38(13):2950–2961, 1999.
- [175] B. W. Pogue, M. S. Patterson, H. Jiang, and K. D. Paulsen. Initial assessment of a simple system for frequency domain diffuse optical tomography. *Physics in Medicine & Biology*, 40(10):1709, 1995.
- [176] S. Posse, E. Ackley, R. Mutihac, J. Rick, M. Shane, C. Murray-Krezan, M. Zaitsev, and O. Speck. Enhancement of temporal resolution and BOLD sensitivity in real-time fMRI using multi-slab echo-volumar imaging. *NeuroImage*, 61(1):115 – 130, 2012.
- [177] P. Poulet, C. V. Zint, M. Torregrossa, W. Uhring, and B. Cunin. Comparison of two time-resolved detectors for diffuse optical tomography: photomultiplier tube–time-correlated single photon counting and multichannel streak camera. In *Optical Tomography and Spectroscopy of Tissue V*, volume 4955 of *Proceedings*, pages 154–163, 2003.
- [178] S. Prahl. Optical absorption of hemoglobin. <https://omlc.org/spectra/hemoglobin/>. Accessed: 2018-09-14.
- [179] J. Prakash, C. Shaw, R. Manjappa, R. Kanhiroan, and P. Yalavarthy. Sparse recovery methods hold promise for diffuse optical tomographic image reconstruction. *IEEE Journal of Selected Topics in Quantum Electronics*, 20(2):74–82, 2013.
- [180] J. W. Prescott. Quantitative imaging biomarkers: The application of advanced image processing and analysis to clinical and preclinical decision making. *Journal of Digital Imaging*, 26(1):97–108, 2013.
- [181] A. Puszka, L. Di Sieno, A. Dalla Mora, A. Pifferi, D. Contini, G. Boso, A. Tosi, L. Hervé, A. Planat-Chrétien, A. Koenig, et al. Time-resolved diffuse optical tomog-



- raphy using fast-gated single-photon avalanche diodes. *Biomedical Optics Express*, 4(8):1351–1365, 2013.
- [182] V. Quaresima, M. Ferrari, A. Torricelli, L. Spinelli, A. Pifferi, and R. Cubeddu. Bilateral prefrontal cortex oxygenation responses to a verbal fluency task: a multichannel time-resolved near-infrared topography study. *Journal of biomedical optics*, 10(1):011012, 2005.
- [183] M. M. Qureshi, J. Brake, H.-J. Jeon, H. Ruan, Y. Liu, A. M. Safi, T. J. Eom, C. Yang, and E. Chung. In vivo study of optical speckle decorrelation time across depths in the mouse brain. *Biomed. Opt. Express*, 8(11):4855–4864, 2017.
- [184] V. R. Ramakrishnan, R. R. Orlandi, M. J. Citardi, T. L. Smith, M. P. Fried, and T. T. Kingdom. The use of image-guided surgery in endoscopic sinus surgery: an evidence-based review with recommendations. *International Forum of Allergy & Rhinology*, 3(3):236–241, 2013.
- [185] R. Re, E. Martinenghi, A. D. Mora, D. Contini, A. Pifferi, and A. Torricelli. Probe-hosted silicon photomultipliers for time-domain functional near-infrared spectroscopy: phantom and in vivo tests. *Neurophotonics*, 3:3 – 3 – 9, 2016.
- [186] J. D. Riley, F. Amyot, T. Pohida, R. Pursley, Y. Ardeshipour, J. M. Kainerstorfer, L. Najafizadeh, V. Chernomordik, P. Smith, J. Smirniotopoulos, et al. A hematoma detector - a practical application of instrumental motion as signal in near infra-red imaging. *Biomed. Opt. Express*, 3(1):192–205, 2012.
- [187] C. S. Robertson, E. L. Zager, R. K. Narayan, N. Handly, A. Sharma, D. F. Hanley, H. Garza, E. Maloney-Wilensky, J. M. Plaum, C. H. Koenig, et al. Clinical evaluation of a portable near-infrared device for detection of traumatic intracranial hematomas. *Journal of Neurotrauma*, 27(9), 2010.
- [188] W. C. Röntgen. On a new kind of rays. *Science*, 3(59):227–231, 1896.
- [189] A. Sakudo. Near-infrared spectroscopy for medical applications: Current status and future perspectives. *Clinica Chimica Acta*, 455:181 – 188, 2016.
- [190] E. Salomatina, B. Jiang, J. Novak, and A. N. Yaroslavsky. Optical properties of normal and cancerous human skin in the visible and near-infrared spectral range. *Journal of Biomedical Optics*, 11(6):064026, 2006.
- [191] J. R. Sanchez, D. Pocci, and M. L. Oelze. A novel coded excitation scheme to improve spatial and contrast resolution of quantitative ultrasound imaging. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(10):2111–2123, 2009.
- [192] M. Schweiger. GPU-accelerated Finite Element Method for Modelling Light Transport in Diffuse Optical Tomography. *Journal of Biomedical Imaging*, 2011:10:10–10:10, 2011.
- [193] M. Schweiger and S. R. Arridge. Application of temporal filters to time resolved data in optical tomography. *Physics in Medicine & Biology*, 44(7):1699, 1999.
- [194] M. Schweiger and S. R. Arridge. The Toast++ software suite for forward and inverse modeling in optical tomography. *Journal of biomedical optics*, 19(4):040801, 2014.

- [195] M. Schweiger, S. R. Arridge, M. Hiraoka, and D. T. Delpy. The finite element method for the propagation of light in scattering media: Boundary and source conditions. *Medical Physics*, 22(11):1779–1792, 1995.
- [196] U. Seibold, B. Wängler, R. Schirmacher, and C. Waengler. Bimodal imaging probes for combined PET and OI: Recent developments and future directions for hybrid agent development. In *BioMed research international*, 2014.
- [197] J. Selb, D. Joseph, and D. A. Boas. Time-gated optical system for depth-resolved functional brain imaging. *Journal of biomedical optics*, 11(4):044008, 2006.
- [198] A. N. Sen, S. P. Gopinath, and C. S. Robertson. Clinical application of near-infrared spectroscopy in patients with traumatic brain injury: a review of the progress of the field. *Neurophotonics*, 3(3), 2016.
- [199] G. F. Serraino and G. J. Murphy. Effects of cerebral near-infrared spectroscopy on the outcome of patients undergoing cardiac surgery: a systematic review of randomised trials. *BMJ Open*, 7(9), 2017.
- [200] Shared Near Infrared File Format Specification. [https://github.com/fangq/snirf/blob/master/snirf\\_specification.md](https://github.com/fangq/snirf/blob/master/snirf_specification.md). Accessed: 2019-07-09.
- [201] C. Shaw and P. Yalavarthy. Effective contrast recovery in rapid dynamic near-infrared diffuse optical tomography using  $l_1$ -norm-based linear image reconstruction method. *Journal of Biomedical Optics*, 17(8):086009, 2012.
- [202] C. R. Simpson, M. Kohl, M. Essenpreis, and M. Cope. Near-infrared optical properties of ex vivo human skin and subcutaneous tissues measured using the Monte Carlo inversion technique. *Physics in Medicine & Biology*, 43(9):2465, 1998.
- [203] J. R. Singer, F. A. Grünbaum, P. Kohn, and J. P. Zubelli. Image reconstruction of the interior of bodies that diffuse radiation. *Science*, 248(4958):990–993, 1990.
- [204] E. Smistad, T. L. Falch, M. Bozorgi, A. C. Elster, and F. Lindseth. Medical image segmentation on GPUs – A comprehensive review. *Medical Image Analysis*, 20(1):1 – 18, 2015.
- [205] R. C. Smith and K. S. Baker. Optical properties of the clearest natural waters (200–800 nm). *Appl. Opt.*, 20(2):177–184, 1981.
- [206] R. Snieder and J. Trampert. Inverse problems in geophysics. In A. Wirgin, editor, *Wavefield Inversion*, pages 119–190, Vienna, 1999. Springer Vienna.
- [207] V. Y. Soloviev. Mesh adaptation technique for Fourier-domain fluorescence lifetime imaging. *Medical Physics*, 33(11):4176–4183, 2006.
- [208] V. Y. Soloviev, J. McGinty, K. B. Tahir, M. A. Neil, A. Sardini, J. V. Hajnal, S. R. Arridge, and F. P. M. Fluorescence lifetime tomography of live cells expressing enhanced green fluorescent protein embedded in a scattering medium exhibiting background autofluorescence. *Opt. Lett.*, 32(14):2034–2036, 2007.
- [209] B. G. Sood, K. McLaughlin, and J. Cortez. Near-infrared spectroscopy: Applications in neonates. *Seminars in Fetal and Neonatal Medicine*, 20(3):164 – 172, 2015.

- [210] J. Steinbrink, H. Wabnitz, H. Obrig, A. Villringer, and H. Rinneberg. Determining changes in NIR absorption using a layered model of the human head. *Physics in Medicine & Biology*, 46(3):879, 2001.
- [211] G. Strangman, J. P. Culver, J. H. Thompson, and D. A. Boas. A Quantitative Comparison of Simultaneous BOLD fMRI and NIRS Recordings during Functional Brain Activation. *NeuroImage*, 17(2):719 – 731, 2002.
- [212] G.-C. Sun, X.-L. Chen, Y.-Z. Hou, X.-G. Yu, X.-D. Ma, G. Liu, L. Liu, J.-S. Zhang, H. Tang, R.-Y. Zhu, D.-B. Zhou, and B.-N. Xu. Image-guided endoscopic surgery for spontaneous supratentorial intracerebral hematoma. *Journal of Neurosurgery*, 127(3):537–542, 2017.
- [213] M. Suzuki, I. Miyai, T. Ono, and K. Kubota. Activities in the frontal cortex and gait performance are modulated by preparation. An fNIRS study. *NeuroImage*, 39(2):600 – 607, 2008.
- [214] M. Taniguchi, T. Yoshimine, D. Cheyne, A. Kato, T. Kihara, H. Ninomiya, M. Hirata, N. Hirabuki, H. Nakamura, and T. Hayakawa. Neuromagnetic fields preceding unilateral movements in dextrals and sinistrals. *Neuroreport*, 9(7):1497–1502, 1998.
- [215] P. Taroni, A. Pifferi, G. Quarto, L. Spinelli, A. Torricelli, F. Abbate, A. M. Villa, N. Balestreri, S. Menna, E. Cassano, and R. Cubeddu. Noninvasive assessment of breast cancer risk using time-resolved diffuse optical spectroscopy. *J. Biomed. Opt.*, 15:15 – 15 – 3, 2010.
- [216] T. Tarvainen, M. Vauhkonen, V. Kolehmainen, and J. P. Kaipio. Hybrid radiative-transfer–diffusion model for optical tomography. *Appl. Opt.*, 44(6):876–886, 2005.
- [217] The European Commission. Cost-benefit analysis for fair research data. policy recommendations – study. *The European Commission*, 2019.
- [218] The European Commission. Cost of not having fair research data. *The European Commission*, 2019.
- [219] M. Tinazzi and G. Zanette. Modulation of ipsilateral motor cortex in man during unimanual finger movements of different complexities. *Neuroscience letters*, 244(3):121–124, 1998.
- [220] A. Torricelli, D. Contini, A. Pifferi, M. Caffini, R. Re, L. Zucchelli, and L. Spinelli. Time domain functional NIRS imaging for human brain mapping. *Neuroimage*, 85:28–50, 2014.
- [221] A. Torricelli, A. Pifferi, P. Taroni, E. Giambattistelli, and R. Cubeddu. In vivo optical characterization of human tissues from 610 to 1010 nm by time-resolved reflectance spectroscopy. *Physics in Medicine & Biology*, 46(8):2227, 2001.
- [222] B. J. Tromberg, A. Cerussi, N. Shah, M. Compton, A. Durkin, D. Hsiang, J. Butler, and R. Mehta. Imaging in breast cancer: Diffuse optics in breast cancer: detecting tumors in pre-menopausal women and monitoring neoadjuvant chemotherapy. *Breast Cancer Research*, 7(6):279, 2005.
- [223] S. Ueda, N. Yoshizawa, T. Shigekawa, H. Takeuchi, H. Ogura, A. Osaki, T. Saeki, Y. Ueda, T. Yamane, I. Kuji, and H. Sakahara. Near-infrared diffuse optical imaging for

- early prediction of breast cancer response to neoadjuvant chemotherapy: A comparative study using 18F-FDG PET/CT. *Journal of Nuclear Medicine*, 57(8):1189–1195, 2016.
- [224] K. Uğurbil, J. Xu, E. J. Auerbach, S. Moeller, A. T. Vu, J. M. Duarte-Carvajalino, C. Lenglet, X. Wu, S. Schmitter, P. F. V. de Moortele, J. Strupp, G. Sapiro, F. D. Martino, D. Wang, N. Harel, M. Garwood, L. Chen, D. A. Feinberg, S. M. Smith, K. L. Miller, S. N. Sotiropoulos, S. Jbabdi, J. L. Andersson, T. E. Behrens, M. F. Glasser, D. C. V. Essen, and E. Yacoub. Pushing spatial and temporal resolution for functional and diffusion MRI in the Human Connectome Project. *NeuroImage*, 80:80 – 104, 2013. Mapping the Connectome.
- [225] R. C. Vannucci and J. M. Perlman. Interventions for perinatal hypoxic–ischemic encephalopathy. *Pediatrics*, 100(6):1004–1114, 1997.
- [226] I. M. Vellekoop. Feedback-based wavefront shaping. *Opt. Express*, 23(9):12189–12206, 2015.
- [227] H. Wabnitz, A. Jelzow, M. Mazurenka, O. Steinkellner, R. Macdonald, D. Milej, N. Żołek, M. Kacprzak, P. Sawosz, R. Maniewski, A. Liebert, S. Magazov, J. Hebden, F. Martelli, P. Di Ninni, G. Zaccanti, A. Torricelli, D. Contini, R. Re, L. Zucchelli, L. Spinelli, R. Cubeddu, and A. Pifferi. Performance assessment of time-domain optical brain imagers, part 2: neuropt protocol. *Journal of Biomedical Optics*, 19:19 – 19 – 12, 2014.
- [228] H. Wabnitz, M. Moeller, A. Liebert, H. Obrig, J. Steinbrink, and R. Macdonald. Time-resolved near-infrared spectroscopy and imaging of the adult human brain. In E. Takahashi and D. F. Bruley, editors, *Oxygen Transport to Tissue XXXI*, pages 143–148. Springer US, 2010.
- [229] H. Wabnitz, D. R. Taubert, M. Mazurenka, O. Steinkellner, A. Jelzow, R. Macdonald, D. Milej, P. Sawosz, M. Kacprzak, A. Liebert, R. Cooper, J. Hebden, A. Pifferi, A. Farina, I. Bargigia, D. Contini, M. Caffini, L. Zucchelli, L. Spinelli, R. Cubeddu, and A. Torricelli. Performance assessment of time-domain optical brain imagers, part 1: basic instrumental performance protocol. *Journal of Biomedical Optics*, 19:19 – 19 – 12, 2014.
- [230] M. Wahl. Time-correlated single photon counting. [https://www.picoquant.com/images/uploads/page/files/7253/technote\\_tcspc.pdf](https://www.picoquant.com/images/uploads/page/files/7253/technote_tcspc.pdf). Accessed: 2018-10-29.
- [231] D. Wang, Y. Wu, and J. Xia. Review on photoacoustic imaging of the brain using nanoprobes. *Neurophotonics*, 3(1):010901, 2016.
- [232] J. Wang and D. Fleischmann. Improving spatial resolution at CT: Development, benefits, and pitfalls. *Radiology*, 289(1):261–262, 2018.
- [233] L. Wang, S. L. Jacques, and L. Zheng. MCML—Monte Carlo modeling of light transport in multi-layered tissues. *Comput. Methods Programs Biomed.*, 47(2):131 – 146, 1995.
- [234] L. V. Wang and H. Wu. *Biomedical optics: Principles and imaging*. Wiley Publisher, 2007.

- [235] W. Weigl, D. Milej, D. Janusek, S. Wojtkiewicz, P. Sawosz, M. Kacprzak, A. Gerega, R. Maniewski, and A. Liebert. Application of optical methods in the monitoring of traumatic brain injury: A review. *Journal of Cerebral Blood Flow & Metabolism*, 36(11):1825–1843, 2016.
- [236] M. N. Wernick, Y. Yang, J. G. Brankov, G. Yourganov, and S. C. Strother. Machine learning in medical imaging. *IEEE Signal Processing Magazine*, 27(4):25–38, 2010.
- [237] M. D. Wilkinson, M. Dumontier, I. J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.-W. Boiten, L. B. da Silva Santos, P. E. Bourne, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3, 2016.
- [238] P. Wintermark, A. Hansen, S. Warfield, D. Dukhovny, and J. Soul. Near-infrared spectroscopy versus magnetic resonance imaging to study brain perfusion in newborns with hypoxic–ischemic encephalopathy treated with hypothermia. *NeuroImage*, 85:287 – 293, 2014.
- [239] S. Wojtkiewicz, T. Durduran, and H. Dehghani. Time-resolved near infrared light propagation using frequency domain superposition. *Biomed. Opt. Express*, 9(1):41–54, 2018.
- [240] M. Wolf, G. Naulaers, F. van Bel, S. Kleiser, and G. Greisen. A review of near infrared spectroscopy for term and preterm newborns. *Journal of Near Infrared Spectroscopy*, 20(1):43–55, 2012.
- [241] H. Xing, W. Bu, Q. Ren, X. Zheng, M. Li, S. Zhang, H. Qu, Z. Wang, Y. Hua, K. Zhao, L. Zhou, W. Peng, and J. Shi. A NaYbF<sub>4</sub>: Tm<sup>3+</sup> nanoprobe for CT and NIR-to-NIR fluorescent bimodal imaging. *Biomaterials*, 33(21):5384 – 5393, 2012.
- [242] C. Xu, H. Vavadi, A. Merkulov, H. Li, M. Erfanzadeh, A. Mostafa, Y. Gong, H. Salehi, S. Tannenbaum, and Q. Zhu. Ultrasound-guided diffuse optical tomography for predicting and monitoring neoadjuvant chemotherapy of breast cancers: Recent progress. *Ultrasonic Imaging*, 38(1):5–18, 2016.
- [243] M. Xu and R. R. Alfano. Fractal mechanisms of light scattering in biological tissue and cells. *Opt. Lett.*, 30(22):3051–3053, 2005.
- [244] Y. Xu, X. Gu, L. L. Fajardo, and H. Jiang. In vivo breast imaging with diffuse optical tomography based on higher-order diffusion equations. *Appl. Opt.*, 42(16):3163–3169, 2003.
- [245] P. Yalavarthy, B. Pogue, H. Dehghani, C. Carpenter, S. Jiang, and K. Paulsen. Structural information within regularization matrices improves near infrared diffuse optical tomography. *Optics Express*, 15(13):8043–8058, 2007.
- [246] P. E. Yoo, S. E. John, S. Farquharson, J. O. Cleary, Y. T. Wong, A. Ng, C. B. Mulcahy, D. B. Grayden, R. J. Ordidge, N. L. Opie, T. J. O'Brien, T. J. Oxley, and B. A. Moffat. 7T-fMRI: Faster temporal resolution yields optimal BOLD sensitivity for functional network imaging specifically at high spatial resolution. *NeuroImage*, 164:214 – 229, 2018. Pushing the spatio-temporal limits of MRI and fMRI.
- [247] H. Yu, J. Park, K. Lee, J. Yoon, K. Kim, S. Lee, and Y. Park. Recent advances in wavefront shaping techniques for biomedical applications. *Current Applied Physics*, 15(5):632 – 641, 2015.

- [248] J. Yu. Symmetric gaussian quadrature formulae for tetrahedral regions. *Computer Methods in Applied Mechanics and Engineering*, 43(3):349 – 353, 1984.
- [249] G. Yuan, U. Alqasemi, A. Chen, Y. Yang, and Q. Zhu. Light-emitting diode-based multiwavelength diffuse optical tomography system guided by ultrasound. *Journal of Biomedical Optics*, 19(12):126003, 2014.
- [250] Zenodo General Policies. <http://about.zenodo.org/policies/>. Accessed: 2019-07-09.
- [251] L. Zhang, T. Cui, and H. Liu. A set of symmetric quadrature rules on triangles and tetrahedra. *Journal of Computational Mathematics*, 27(1):89–96, 2009.
- [252] R. Zhang, H. Wang, R. Tewari, G. Schmidt, and D. Kakrania. Big Data for medical image analysis: A performance study. In *2016 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, pages 1660–1664, 2016.
- [253] X. Zhang. Instrumentation in diffuse optical imaging. *Photonics*, 1(1):9–32, 2014.
- [254] Q. Zhao, L. Ji, and T. Jiang. Improving depth resolution of diffuse optical tomography with a layer-based sigmoid adjustment method. *Optics express*, 15(7):4018–4029, 2007.
- [255] Q. Zhu, S. Tannenbaum, and S. H. Kurtzman. Optical tomography with ultrasound localization for breast cancer diagnosis and treatment monitoring. *Surgical Oncology Clinics of North America*, 16(2):307 – 321, 2007.
- [256] G. Zonios and A. Dimou. Light scattering spectroscopy of human skin in vivo. *Opt. Express*, 17(3):1256–1267, 2009.
- [257] J. Zouaoui, L. D. Sieno, L. Hervé, A. Pifferi, A. Farina, A. D. Mora, J. Derouard, and J.-M. Dinten. Chromophore decomposition in multispectral time-resolved diffuse optical tomography. *Biomed. Opt. Express*, 8(10):4772–4787, 2017.



# Glossary

- APD** Avalanche Photodiode
- BOLD** Blood-Oxygen-Level-Dependent Imaging
- CDD** Charge-Coupled Device
- CT** Computed Tomography
- CW** Continuous-Wave System
- DA** Diffusion Approximation
- DTOF** Distribution of photons Times-of-Flight
- FD** Frequency Domain System
- FEM** Finite Element Method
- FT** Fourier Transform
- GPU** Graphics Processing Unit
- HHb** Deoxygenated Hemoglobin
- MRI** Magnetic Resonance Imaging
- NIRS** Near-Infrared Spectroscopy
- O2Hb** Oxygenated Hemoglobin
- PD** Solid-state Photodiode
- PDE** Partial Derivative Equation
- PMT** Photomultiplier Tube
- RTE** Radiative Transfer Equation
- SiPM** Silicon Photomultiplier Detector
- SNR** Signal-to-Noise Ratio
- SPAD** Single Photon Avalanche Diode
- TCSPC** Time-Correlated Single Photon Counting



**tHb** Total Hemoglobin

**TR** Time-Resolved System

**tr-DOT** Time-Resolved Diffuse Optical Tomography

**TPSF** Temporal Point Spread Function

## Résumé en français

La surveillance non invasive des patients est essentielle pour l'avenir des dispositifs d'imagerie médicale. Dans ce contexte, la biophotonique est un domaine qui suscite beaucoup d'intérêt dans le domaine médical car sa technologie est intrinsèquement non invasive, miniaturisable, portable, sûre et abordable. Pour ces raisons, la Commission Européenne (CE) a récemment choisi la photonique et ses applications biomédicales comme l'une des technologies clés et d'avenir permettant d'améliorer la compétitivité de l'Union européenne sur les marchés [218].

Dans ce contexte, des laboratoires européens intensifient leur travail dans les différents domaines de la biophotonique. La spectroscopie d'optique diffuse et l'imagerie basée sur la lumière proche infrarouge sont l'un de ces domaines de recherche prometteurs. Pour étudier les applications de l'optique diffuse pour le suivi des paramètres neurologiques, la CE a apporté un soutien financier au réseau de formation initiale BitMap (Brain injury and trauma monitoring using advanced photonics) [137]. L'idée centrale de ce réseau est de rassembler des équipes, des chercheurs des doctorants, de divers horizons (électronique, photonique, mathématiques et essais cliniques), de les faire collaborer pour relever les principaux défis technologiques et cliniques des soins d'urgence neurologique à base de photonique. Par ce travail de thèse j'ai fait partie de ce réseau et contribué à améliorer la reconstruction tomographique basée sur la technologie optique diffuse.

Cette technique de tomographie optique diffuse dans le proche infrarouge (DOT : Diffuse Optical Tomography) permet de sonder de manière non invasive et en profondeur le tissu humain en reconstituant les paramètres de la composition des tissus biologiques tels que les concentrations d'hémoglobine oxygénée et désoxygénée dans le sang.

Dans ce manuscrit de thèse, je décris les nouvelles améliorations que j'ai développées dans le domaine des algorithmes DOT à résolution temporelle. Après un premier chapitre d'introduction et de positionnement, dans le deuxième chapitre, je détaille les bases de l'imagerie dans le proche infrarouge et l'état actuel des techniques de tomographie optique diffuse. Dans le chapitre suivant, je présente une nouvelle méthode de calcul des «datatypes» pour la reconstruction tomographique de mesures résolues dans le temps. Les résultats montrent qu'avec cette nouvelle approche, le bruit des «datatypes» est décorrélé et la résolution en profondeur des reconstructions est améliorée de manière significative par rapport aux algorithmes de l'état de l'art. Dans le quatrième chapitre, je me concentre sur la régularisation du problème de tomographie et je décris deux approches différentes pour effectuer la régularisation de la variation totale sur des maillages irréguliers. Les al-

algorithmes développés sont le résultat d'une collaboration avec l'Université de Birmingham. Le cinquième chapitre est consacré à la vérification des résultats théoriques et numériques développées précédemment. Des mesures expérimentales sur des fantômes solides et des expériences in vivo sur des sujets humains ont été réalisées. En collaboration avec Politecnico di Milano, j'ai testé un système tomographique à double longueur d'onde à résolution temporelle avec deux sondes de trois sources et quatre détecteurs de photomultiplicateurs en silicium. Des expériences réelles sur sujet humain, d'occlusion de bras ont été effectuées pour valider le système. Ensuite, j'ai effectué des expériences d'activation du cortex moteur basées sur le tapotement des doigts sur trois sujets adultes différents. Les résultats montrent qu'il est possible de surveiller avec une résolution temporelle d'une seconde l'activation du cortex moteur sur les deux hémisphères et que les informations spatiales et en profondeur peuvent également être récupérées. Enfin, au sixième chapitre, je présente au lecteur les efforts déployés par le réseau BitMap pour promouvoir la normalisation du domaine de l'optique diffuse. Je décris le travail que j'ai effectué pour construire un jeu de données ouvert avec les mesures effectuées sur 28 instruments de huit institutions européennes différentes à l'aide de trois protocoles européens validés.

Ci-après, je résume les résultats présentés dans cette thèse.

## Chapitre 2. Contexte

Tous les jours, dans les hôpitaux et les centres médicaux, une large gamme de technologies d'imagerie médicale est utilisée: d'une simple radiographie à rayons X à une IRM fonctionnelle plus sophistiquée. Toutes ces technologies sont basées sur différents phénomènes physiques et ont des propriétés et des caractéristiques différentes qui les rendent appropriées en fonction de la situation clinique. Par exemple, la résolution spatiale, par contraste et temporelle peut aller de plusieurs ordres de grandeur. De plus, le coût, les contre-indications et la complexité du système sont également des caractéristiques essentielles pour déterminer le succès d'une technologie dans les cliniques. La lumière est l'un de ces phénomènes physiques largement utilisés dans les environnements cliniques. Par exemple, la tomographie par cohérence optique est l'une de ces réussites. Les principaux avantages des systèmes médicaux à base optique sont leur évolutivité, leur portabilité et leur flexibilité. Souvent, ils peuvent être combinés avec d'autres technologies d'imagerie médicale, leur coût est faible comparé aux technologies d'imagerie actuelles et est assez sûr. En outre, les technologies basées sur l'optique peuvent potentiellement obtenir des informations chimiques à partir de l'interaction des photons avec les tissus.

En ce qui concerne la résolution spatiale, la lumière peut pénétrer dans le tissu humain sur de grandes distances (de l'ordre de quelques mètres). Cependant, il est connu que la probabilité qu'un tel événement se produise est très faible et que le bruit masquerait toute information que ces photons pourraient véhiculer. Autrement dit, la nature très diffuse du tissu humain limite la profondeur moyenne de pénétration à quelques centimètres dans le meilleur des cas. Le corps humain composé de plusieurs chromophores contient évidemment beaucoup d'hémoglobine qui est l'un des chromophores les plus importants. Il se situe au niveau des globules rouges et sa tâche principale est de transporter l'oxygène dans le corps humain. L'hémoglobine existe dans deux états: oxygéné ou désoxygéné; dans le premier cas, l'hémoglobine contient des molécules d'oxygène et est connue sous le nom d'oxyhémoglobine ( $O_2Hb$ ); dans le dernier cas, l'hémoglobine est exempte d'oxygène et est

appelée désoxyhémoglobine (HHb). Ce chromophore dans ces deux états est très visible. Plusieurs études ont été publiées qui caractérisent l'absorption de chaque état d'hémoglobine (voir Figure \*1). L'absorption de l'hémoglobine à des longueurs d'onde courtes ( $< 600$  nm) est importante. Néanmoins, entre 600 et 900 nm, l'absorption diminue considérablement et la différence entre l'oxy- et la désoxy-hémoglobine est évidente; cette plage est généralement appelée fenêtre optique ou fenêtre thérapeutique. Dans cette plage, la lumière n'est pas très absorbée et la séparation permet de déduire la concentration de chaque état d'hémoglobine en utilisant la loi de Beer-Lambert.

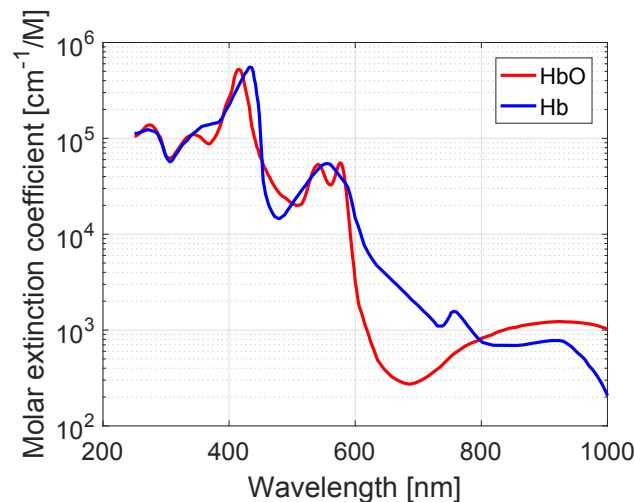


Figure \*1: Coefficient d'extinction molaire de l'oxy- et de la désoxyhémoglobine dans le spectre visible. Données extraites de [178].

Au cours des vingt dernières années, la technologie optique diffuse a évolué afin d'améliorer la sensibilité aux (1) couches plus profondes du tissu et (2) aux changements d'absorption. Dans cette thèse, seuls les systèmes résolus en temps ont été considérés. L'idée centrale de ces systèmes est de mesurer le temps de vol de la distribution des photons (DTOF), car le temps d'arrivée des photons peut être associé de manière probabiliste à une profondeur de pénétration dans le milieu et, par conséquent, permet de coder l'information de profondeur.

Pour résoudre ce problème de reconstruction, la tomographie optique diffuse fait appel à des algorithmes et méthodes de calcul. Ils sont composés de deux parties: les modèles directs et inverses. Les modèles directs simulent le comportement photonique au niveau du tissu humain. Ces modèles doivent être précis et efficaces car ils constituent une partie essentielle de l'algorithme de reconstruction. Dans la littérature, deux modèles principaux sont utilisés: l'équation de transfert radiatif et l'approximation de diffusion. La première est une équation intégral-différentielle qui est résolue via les méthodes de Monte-Carlo. La deuxième est une équation différentielle partielle dépendant du temps qui peut être résolue par les méthodes des éléments finis ou des volumes finis. La deuxième partie d'un algorithme de tomographie diffuse est le modèle inverse. Dans ce travail, j'ai utilisé l'approximation de Born pour retrouver les propriétés optiques du milieu. Un schéma d'un algorithme typique est présenté à la Figure \*2.

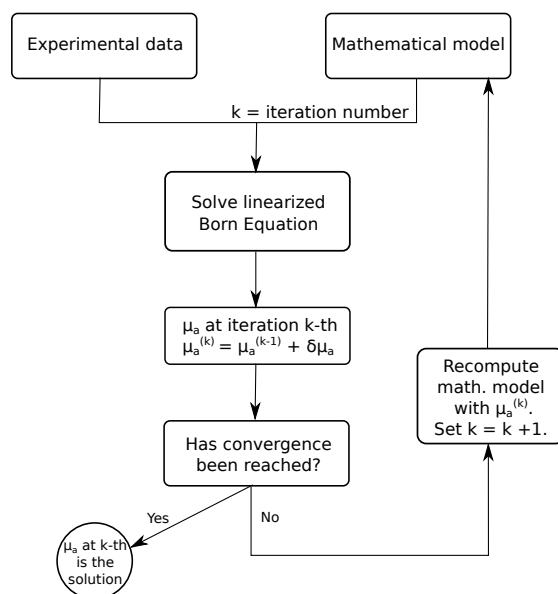


Figure \*2: Algorithme de reconstruction par absorption basé sur l'équation de Born linéarisée.

Dans le chapitre suivant, je décris une nouvelle méthode de calcul des «datatypes» pour effectuer une reconstruction tomographique. Cette méthode améliorera la résolution en profondeur et la quantification d'absorption des reconstructions tomographiques.

### Chapitre 3. Etude de «datatypes» pour la tomographie optique diffuse à résolution temporelle

Le modèle direct est un élément clé du processus de reconstruction. L'équation d'approximation de diffusion en fonction du temps est le modèle le plus largement utilisé sur le terrain car elle peut être résolue à l'aide de la méthode des éléments finis. Cependant, le processus est encore lent, car le temps doit être discrétisé en petites étapes afin de garantir la stabilité et la convergence. Pour cette raison, plusieurs auteurs ont proposé d'utiliser des «datatypes» de fenêtres temporelles de la forme  $w(t) = t^n e^{-pt}$  car ils sont plus rapides à calculer que le débit de fluence lui-même.

Les «datatypes» de la forme  $w(t) = t^n e^{-pt}$ , bien que rapides à calculer, sont très corrélés. Pour cette raison, une des priorités dans mon travail a été de proposer de calculer des «datatypes» en utilisant le théorème de Plancherel,

$$\Gamma = \int_{-\infty}^{\infty} u(t)\overline{w(t)} dt = \int_{-\infty}^{\infty} U(f)\overline{W(f)} df,$$

calculant les «datatypes» dans le domaine fréquentiel. Les avantages de cette méthode sont que les «datatypes» ne sont pas limités à la forme  $w(t) = t^n e^{-pt}$  et qu'un plus grand ensemble de fonctions peut être utilisé. La Figure \*3 ci dessous montre les matrices de corrélation pour les fenêtres de Mellin-Laplace,  $w(t) = t^n e^{-pt}$ , gaussienne et Tukey. Les fenêtres de Mellin-Laplace sont de plus en plus corrélées à mesure que l'ordre augmente.

Toutefois, les fenêtres gaussienne et Tukey ne se chevauchent pas et, par conséquent, ne sont pas corrélées.

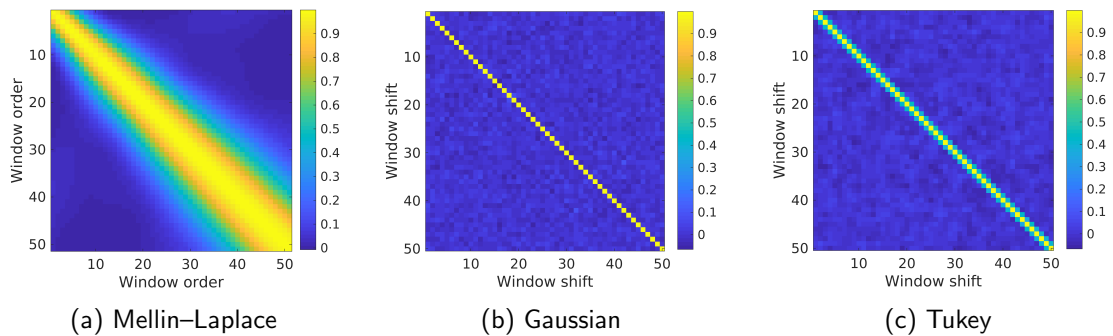


Figure \*3: Matrices de corrélation de Mellin-Laplace (50 premières ordres,  $p = 7$ ), Gaussienne ( $t_0 = 0.16$  ns décalés,  $\sigma = 0.05$ ) et de fenêtres Tukey ( $t_0 = 0.16$  ns décalés,  $\alpha = 0.25$  et  $t^* = 0.2$  ns).

Les avantages de ces nouveaux «datatypes» sont illustrés à la Figure \*4 ci dessous; on voit qu'avec les «datatypes» Mellin-Laplace, il est impossible de localiser une inclusion profonde ( $> 3.0$  cm). Par contre, en utilisant les fenêtres Gaussienne ou de type Tukey, on est capable de déterminer avec précision la profondeur de la sphère. De plus, la quantification d'absorption est bien meilleure qu'avec les «datatypes» Mellin-Laplace.

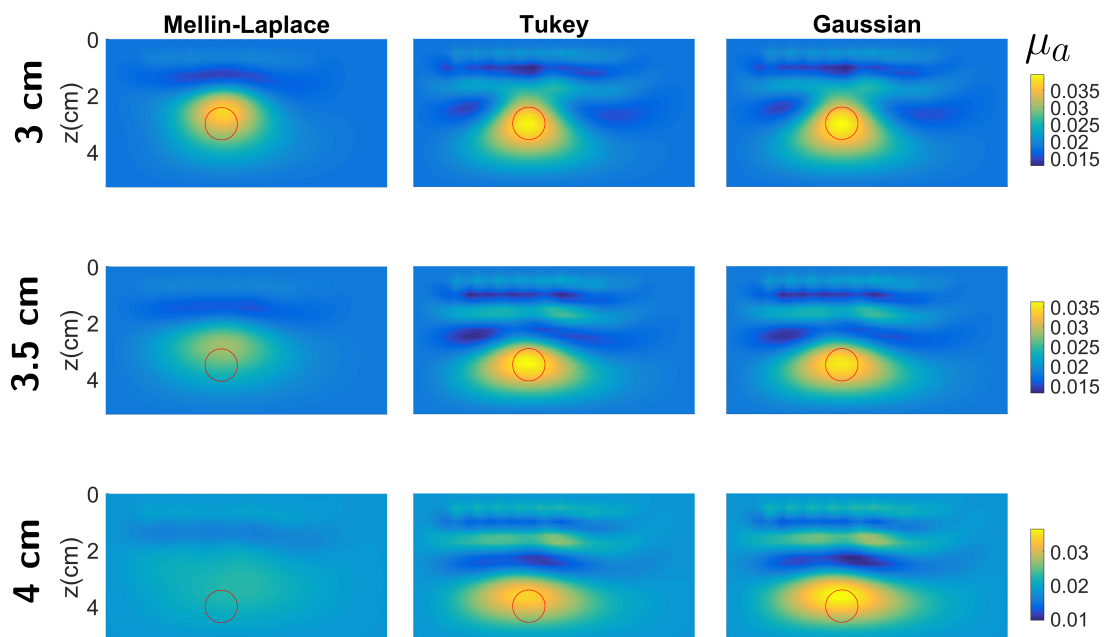


Figure \*4: Absorption,  $\mu_a$ , reconstructions utilisant des fenêtres de Mellin - Laplace, Tukey et Gaussienne pour des inclusions de 3 cm à 4 cm deep. Le cercle rouge indique l'emplacement correct de l'inclusion.

Par conséquent, ces nouveaux «datatypes» n'améliorent pas seulement la reconstruction dans les couches profondes des tissus. Mais sont également plus efficaces que la résolution directe

de l'approximation de diffusion en fonction du temps. Il pourrait être possible à l'avenir de construire un système qui calcule directement les fenêtres des signaux optiques à la place des courbes résolues en temps. Ce système hypothétique pourrait être potentiellement plus rapide que les systèmes actuels sans compromettre la précision tomographique.

## Chapitre 4. Régularisation à variation totale pour tomographie optique diffuse

Le modèle inverse issu du DOT est connu pour être un problème mal posé. La principale caractéristique des problèmes mal posés est qu'ils sont très sensibles aux perturbations des conditions initiales. Pour cette raison, les solutions aux problèmes mal posés sont généralement régularisées en ajoutant certaines exigences de lissage. La variation totale (TV) est l'une des techniques de régularisation les plus utilisées pour la reconstruction d'images. Cependant, ses performances sur les domaines complexes anatomiques n'ont pas été analysées dans le contexte de la DOT. La régularisation de la variation totale minimise la norme  $l_1$  du gradient, ce qui favorise des solutions constantes par morceaux. L'opérateur de gradient est bien défini dans un support continu. Cependant, pour les domaines discrétisés, l'opérateur de gradient doit être redéfini.

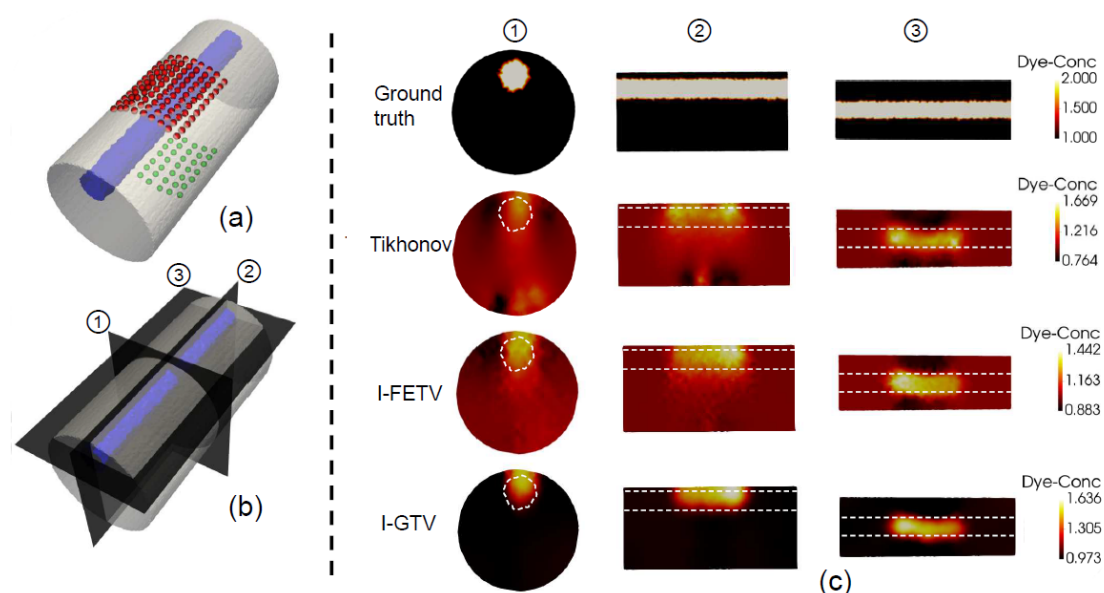


Figure \*5: (a) Géométrie cylindrique avec sources et détecteurs. (b) Les différents plans de vue. (c) De haut en bas: vérité au sol, Tikhonov, I-FETV et I-GTV. Les colonnes représentent chaque plan de vue. (Figure tirée de [135]).

En collaboration avec l'université de Birmingham, j'ai développé deux approches différentes pour définir l'opérateur de gradient pour les maillages non structurés: les approches à éléments finis (FETV) et à base de graphes (GTV). Dans le FETV, la géométrie est discrétisée en utilisant des éléments triangulaires et tétraédriques pour les géométries bidimensionnelles et tridimensionnelles, respectivement. Ces éléments sont composés d'un ensemble de sommets et d'arêtes et ne se chevauchent pas. Cependant, en GTV, le domaine est discrétisé par un graphe pondéré dans lequel chaque nœud peut être connecté au reste des nœuds.

Les simulations numériques et les fantômes expérimentaux montrent que FETV offre de meilleures performances pour les maillages haute résolution. Toutefois, lorsque la résolution du maillage est faible (par exemple, pour des maillages tridimensionnels de grande taille), GTV a des artefacts plus faibles aux bords du maillage. La Figure \*5 montre une comparaison entre la régularisation de Tikhonov (classique) et les techniques développées à la pointe de la technologie. On peut voir que la régularisation de Tikhonov est sujette aux artefacts proches des bords. Au contraire, les approches GTV et FETV ne présentent pas ces artefacts. On remarque aussi qu'avec l'approche GTV, on obtient une solution plus homogène qu'avec l'approche FETV.

## Chapitre 5. Expériences in vivo avec un système optique diffus résolu en temps

La prochaine étape de mon travail consistait à tester les algorithmes de tomographie optique diffuse développés sur des données expérimentales. Pour cette raison, j'ai collaboré avec Politecnico di Milano qui a proposé d'effectuer des mesures avec un système résolu en temps et à double longueur d'onde (670 et 820 nm) fourni par deux sondes à trois fibres et quatre détecteurs photomultiplicateurs au silicium.

La première partie de ce travail a consisté à tester le système optique et les algorithmes dans un fantôme solide avec une inclusion à une profondeur de 1.5 cm. Ensuite, des expériences d'occlusion de bras ont été effectuées chez des sujets adultes afin de valider le système. Les expériences d'occlusion sont faciles à réaliser et le processus physiologique qui les sous-tend est bien connu. Deux types d'expériences d'occlusion de bras ont été réalisées: occlusions veineuse et artérielle. La différence réside dans la pression exercée sur le brassard et ses conséquences physiologiques. Pour l'occlusion veineuse, lorsque la pression du brassard ne dépasse pas la pression systolique, seules les veines sont bloquées. Par conséquent, le flux sanguin de sortie du bras est bloqué pendant que les artères sont encore ouvertes ; on observe l'augmentation à la fois de l'hémoglobine oxygénée et désoxygénée dans le bras lors de l'occlusion, et que la quantité d'oxygène hémoglobineux est supérieure à celle désoxygénée puisque le taux de consommation de sang oxygéné par le bras est inférieur au débit sanguin oxygéné des artères. Pour les occlusions artérielles, la pression du brassard est au moins deux fois supérieure à la pression systolique des sujets. Dans ce cas, les veines et les artères sont bloquées. L'entrée et la sortie de sang sont donc totalement bloquées au niveau du bras. Dans l'ensemble, la quantité d'hémoglobine oxygénée diminuera car elle est consommée par les muscles des bras et elle sera transformée en hémoglobine désoxygénée. Pour les trois sujets, les résultats physiologiques étaient tels que prévus.

Après la validation avec des tests d'occlusion du bras, des expériences d'activation du cortex moteur ont été réalisées à l'aide d'un test de tapotement des doigts. Le protocole était le suivant: repos de 20 s – tapotement de doigts de 20 s – repos de 20 s, qui était répété cinq fois par sujet. La séquence du tapotement des doigts effectuée était la suivante: (1) le pouce touche l'index, (2) le pouce touche l'annulaire, (3) le pouce touche le majeur et (4) le pouce touche l'auriculaire.

La Figure \*6 montre l'activation au niveau de l'hémisphère controlatérale lors du tapotement de la main gauche pour l'un des sujets. L'activation est décrite par une augmentation de l'hémoglobine oxygénée et une légère diminution de l'hémoglobine désoxygénée



(généralement un tiers de l'hémoglobine oxygénée). Ces résultats confirment qu'il est possible de surveiller l'activité du cortex cérébral chez des sujets adultes avec une résolution inférieure à une seconde en utilisant une technologie à résolution temporelle avec des détecteurs SiPM.

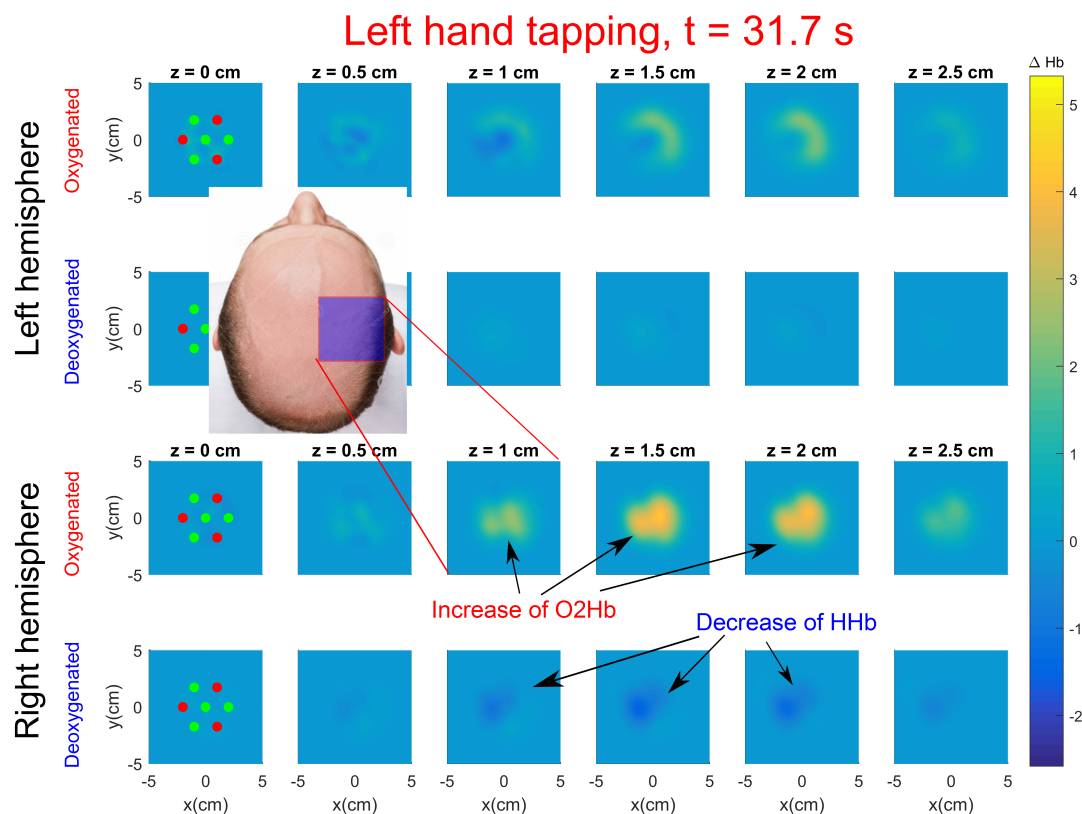


Figure \*6: Reconstruction tomographique de l'activation de l'hémisphère droit lors d'une expérience de tapotement à la main gauche réalisée par un sujet adulte.

## Chapitre 6. Le BitMap Open Data

Le mouvement Open Science a gagné en popularité au cours des dernières années. Son objectif est de rendre la recherche scientifique ouverte à toute la société en publiant non seulement les résultats, mais également les données et les logiciels utilisés au cours de la recherche. Dans les communautés de spectroscopie optique diffuse et d'imagerie, l'approche Open Data n'est pas encore complètement développée. Les données provenant d'expériences sont dispersées sur Internet, difficiles à comparer en raison du manque de métadonnées / d'informations documentées ou tout simplement indisponibles. De plus, le manque de données publiques sur les fantômes standards rend également la normalisation des instruments d'optique diffuse beaucoup plus difficile à réaliser. Par conséquent, un référentiel avec des données optiques diffuses documentées, conformes à des normes bien définies, fait toujours défaut.

| ID | Système                              | Labo      | Technique | Application  | TRL |
|----|--------------------------------------|-----------|-----------|--------------|-----|
| 1  | Clinical broadband spectrometer      | POLIMI    | TR        | Spectromètre | 5   |
| 2  | 3mm SiPM                             | POLIMI    | TR        | Oxymètre     | 3   |
| 3  | HPM based spectrometer               | PTB       | TR        | Spectromètre | 4   |
| 4  | NIRO 200NX                           | UHB/UoB   | CW        | Oxymètre     | 8   |
| 5  | ISS OXIPLEX-TS                       | UHB/UoB   | FD        | Oxymètre     | 8   |
| 6  | ISS IMAGENT                          | UHB/UoB   | FD        | Oxymètre     | 8   |
| 7  | Multiwavelength system               | IBIB      | TR        | Spectromètre | 5/6 |
| 8  | TD-DCS system                        | IBIB      | TR        | Perfusion    | 4   |
| 9  | SRS based CW system                  | UCL       | CW        | Spectromètre | 6   |
| 10 | TRS-DCS FLOWer                       | ICFO/VdH  | TR        | Oxymètre     | 7   |
| 11 | MCP based system                     | PTB       | TR        | Spectromètre | 4   |
| 12 | Fetus clinical instrument, HC1       | ICFO/VdH  | TR        | Oxymètre     | 6   |
| 13 | Clinical TD oximeter                 | IBIB      | TR        | Oxymètre     | 6   |
| 14 | 8 channel system                     | IBIB      | TR        | Oxymètre     | 6   |
| 15 | MAESTROS                             | UCL       | TR        | Spectromètre | 4   |
| 16 | LUCA                                 | POLIMI    | TR        | Spectromètre | 6   |
| 17 | TD oximeter                          | POLIMI    | TR        | Oxymètre     | 6   |
| 18 | clinical TD oximeter                 | POLIMI    | TR        | Oxymètre     | 6   |
| 19 | wearable TD device                   | POLIMI    | TR        | Spectromètre | 5   |
| 20 | Artinis                              | POLIMI    | CW        | Oxymètre     | 8   |
| 21 | Mammograph                           | POLIMI    | TR        | Imagerie     | 6   |
| 22 | Multiwavelength "Fruit" spectrometer | POLIMI    | TR        | Spectromètre | 4   |
| 23 | OCTOPUS                              | POLIMI    | TR        | Spectromètre | 4   |
| 24 | Laboratory Spectromètre System       | POLIMI    | TR        | Spectromètre | 6   |
| 25 | Laboratory broadband spectrometer    | POLIMI    | TR        | Spectromètre | 6   |
| 26 | Lab TD-DCS                           | POLIMI    | TR        | Spectromètre | 4   |
| 27 | Mammograph MH                        | POLIMI    | TR        | Imagerie     | 6   |
| 28 | SFD Imager                           | ICube/UoS | FD        | Imagerie     | 4   |
| 29 | APD based Spectromètre               | ICube/UoS | TR        | Spectromètre | 4   |

Table \*1: Caractéristiques des instruments utilisés pendant la campagne. TRL = Niveau de préparation technologique du système testé (par exemple, TRL3 = composant / sous-système, TRL4 = prototype de laboratoire, TRL5 = prototype clinique, TRL6 = système clinique déjà mis en évidence dans des cliniques, TRL8 = dispositif clinique commercial).

Dans cette thèse, j'ai présenté l'ensemble de données obtenu lors d'une campagne d'évaluation des performances impliquant plusieurs instruments européens d'optique diffuse. L'idée de la campagne était double : premièrement, dresser un tableau de l'état actuel des instruments européens en optique diffuse, c'est-à-dire illustrer les différents types de systèmes utilisés et décrire la variabilité des résultats. Deuxièmement, créer un ensemble ouvert de données afin que les chercheurs travaillant dans le traitement de données puissent tester différentes techniques permettant de récupérer les propriétés optiques. La campagne a été réalisée par un chercheur du réseau BitMap. Ce chercheur s'est rendu dans 10 institutions et a effectué des mesures à l'aide de 29 instruments différents (voir la Table \*1) en appliquant trois protocoles internationalement reconnus (protocole BIP, MEDPHOT et nEUROPt).

Mon travail a consisté à collecter toutes les données brutes et à les organiser dans un format

contenant suffisamment de métadonnées et les publier sur l'internet. Le format de données utilisé était la spécification de format SNIRF (Shared Near Infrared File), basée sur le format HDF5.

Je termine ce travail en proposant une conclusion qui reprend l'ensemble des travaux et des perspectives d'amélioration et de poursuite d'études.

## Abstract

Noninvasive and continuous monitoring of patients are key features in the future of medical imaging. Biophotonics is a field that is attracting a lot of interest because its technology is intrinsically noninvasive and potentially miniaturizable and wearable. Regarding the imaging of human tissue using photonics, it has been proven that near-infrared diffuse optical tomography (DOT) permits to probe noninvasively and in depth the human tissue by reconstructing parameters of the composition of biological tissues such as the concentrations of oxygenated and deoxygenated hemoglobin in the blood. In this thesis, I describe the novel improvements I developed in the field of time-resolved DOT algorithms. First, I introduce a novel method to compute datatypes for tomographic reconstruction of time-resolved measurements. The results show that with this new approach the noise of datatypes are decorrelated and resolution in depth of reconstructions is improved significantly for inclusions deeper than 2.5 centimeters. After, I describe two different approaches to perform total variation regularization for DOT reconstruction on irregular meshes. The knowledge developed in previous parts was applied to in-vivo experiments on human subjects. In collaboration with Politecnico di Milano, I tested a time-resolved tomographic system with two probes of three source fibers and four silicon photomultiplier detectors each. Arm occlusion experiments were performed to validate the technology. After, I did motor cortex activation experiments on three different subjects. The results show that it is possible to monitor with one-second resolution the motor cortex activation and that spatial and in depth information can also be retrieved. Finally, I introduce the reader to the effort that is being done at BitMap network to push the standardization of diffuse optics field. I describe the work I did to build an open dataset with the measurements performed at twenty-eight instruments from eight different European institutions using three validated European protocols.

**Keywords:** diffuse optical tomography, inverse problem, functional near infrared spectroscopy

## Résumé

La surveillance non invasive et continue des patients est un élément clé de l'avenir de l'imagerie médicale. La biophotonique est un domaine qui suscite beaucoup d'intérêt, car sa technologie est intrinsèquement non invasive, potentiellement miniaturisable et portable. En ce qui concerne l'imagerie de tissus humains en photonique, il a été prouvé que la tomographie optique diffuse dans le proche infrarouge (DOT) permettait de sonder de manière non invasive et en profondeur les tissus humains en reconstruisant des paramètres de la composition des tissus biologiques, tels que les concentrations en hémoglobine oxygénée et désoxygénée du sang. Dans cette thèse, je décris les nouvelles améliorations que j'ai développées dans le domaine des algorithmes DOT sur des acquisitions résolues en temps. Premièrement, je présente une nouvelle méthode de calcul des types de données pour la reconstruction tomographique de mesures résolues en temps. Les résultats montrent qu'avec cette nouvelle approche, le bruit des nouveaux types de données est décorrélié et la résolution en profondeur des reconstructions est améliorée de manière significative pour les inclusions plus profondes que 2,5 centimètres. Ensuite, je décris deux approches différentes pour effectuer la régularisation «variation totale» pour la reconstruction DOT sur des maillages irréguliers. Les connaissances développées dans les parties précédentes ont été appliquées à des expériences in vivo sur des sujets humains. En collaboration avec Politecnico di Milano, j'ai testé un système tomographique à résolution temporelle avec deux sondes de trois sources et quatre détecteurs (photomultiplicateurs en silicium). Des expériences d'occlusion de bras ont été réalisées pour valider la technologie. Après, j'ai fait des expériences d'activation du cortex moteur sur trois sujets différents. Les résultats montrent qu'il est possible de surveiller l'activation du cortex moteur avec une résolution d'une seconde et que des informations spatiales et en profondeur peuvent également être récupérées. Enfin, je présente au lecteur les efforts déployés par le réseau BitMap pour pousser la normalisation du domaine de l'optique diffuse. Je décris le travail que j'ai effectué pour construire un jeu de données public avec les mesures effectuées sur 28 instruments de huit institutions européennes différentes à l'aide de trois protocoles européens validés.

**Mots-clés:** tomographie optique diffuse, problème inverse, spectroscopie fonctionnelle proche infrarouge