IMPROVED MODELING AND IMAGE GENERATION FOR FLUORESCENCE MOLECULAR TOMOGRAPHY (FMT) AND POSITRON EMISSION TOMOGRAPHY (PET)

by
Yansong Zhu

A dissertation submitted to Johns Hopkins University in conformity
with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
September 2020

© 2020 Yansong Zhu
All rights reserved

Abstract

In this thesis, we aim to improve quantitative medical imaging with advanced image generation algorithms. We focus on two specific imaging modalities: fluorescence molecular tomography (FMT) and positron emission tomography (PET).

In the case of FMT, we present a novel photon propagation model for its forward model, and in addition, we propose and investigate a reconstruction algorithm for its inverse problem. In the first part, we develop a novel Neumann-series-based radiative transfer equation (RTE) that incorporates reflection boundary conditions in the model. In addition, we propose a novel reconstruction technique for diffuse optical imaging that incorporates this Neumann-series-based RTE as forward model. The proposed model is assessed using a simulated 3D diffuse optical imaging setup, and the results demonstrate the importance of considering photon reflection at boundaries when performing photon propagation modeling. In the second part, we propose a statistical reconstruction algorithm for FMT. The algorithm is based on sparsity-initialized maximum-likelihood expectation maximization (MLEM), taking into account the Poisson nature of data in FMT and the sparse nature of images. The proposed method is compared with a pure sparse reconstruction method as well as a uniform-initialized MLEM reconstruction method. Results indicate the proposed method is more robust to noise and shows improved qualitative and quantitative performance.

For PET, we present an MRI-guided partial volume correction algorithm for brain imaging, aiming to recover qualitative and quantitative loss due to the limited resolution of PET system, while keeping image noise at a low level. The proposed method is based on an iterative deconvolution model with regularization using parallel level sets. A non-smooth optimization algorithm is developed so that the proposed method can be feasibly applied for 3D images and avoid additional blurring caused by conventional smooth optimization process. We evaluate the proposed method using both simulation data and *in vivo* human data collected from the Baltimore Longitudinal Study of Aging (BLSA). Our proposed method is shown to generate images with reduced noise and improved structure details, as well as increased number of statistically significant voxels in study of aging. Results demonstrate our method has promise to provide superior performance in clinical imaging scenarios.

Thesis Committee

Prof. Arman Rahmim
Department of Electrical and Computer Engineering,
Department of Radiology and Radiological Sciences (advisor, primary reader)

Prof. Yong Du
Department of Radiology and Radiological Sciences (secondary reader)

Prof. Jin Kang
Department of Electrical and Computer Engineering

Prof. Trac Tran
Department of Electrical and Computer Engineering

Acknowledgements

During the past five years, I have received help and support from different people. I would like to express my gratitude to my advisor, Dr. Arman Rahmim for all the support, guidance, and freedom to explore ideas that he provided during my entire PhD. He taught me how to become a better researcher from different aspects. He inspired me to dig deeper into my ideas, encouraged me to reach out to different people, and shared his thoughts on time management and goal setting. It has been a great fortune for me to study under his supervision.

I would like to thank Dr. Abhinav Jha for his guidance and help, especially during my early years. He and Arman together led me into the world of research. He provided lots of useful tools and references when I almost knew nothing about research. He had always been patient discussing with me on different problems despite his busy schedule. His hardwork also motivated me a lot.

I would like to thank Dr. Yong Du, for agreeing to be in my dissertation committee as the reader. I would like to thank Dr. Jin Kang, for providing helpful discussions and agreeing to be in my PhD qualifying exam, graduate board oral exam, and my dissertation committee. I would like to thank Dr. Trac Tran for being a member of my dissertation committee and for teaching an interesting compressed sensing course where I had learnt a lot. I would like to thank Dr. Xingde Li, Dr. Eric Frey, and Dr. Web Stayman for being in my graduate board oral exam committee. I would like to thank Dr. Jerry Prince, for agreeing to be in my PhD qualifying exam committee, for providing me the opportunity of working as a teaching assistant in his course,

and for providing a wonderful medical imaging systems course where I first glimpsed into the world of medical imaging. I would like to thank people from partial volume correction meeting group. I would like to thank Dr. Murat Bilgel. He has provided me lots of useful suggestions and has always been willing to try my new algorithms on clinical datasets. I would like to thank Dr. Susan Resnick, Dr. Dean F. Wong, and Dr. Oliver Rousset for providing helpful discussions. I would like to acknowledge support from NIH BRAIN Initiative Award (R24 MH106083), NIH R21 grant AG056142, BC Cancer Foundation and SNMMI 2020 Bradley-Alavi fellowship. This work would not have taken place without these supports.

I would like to express my gratitude to members from quantitative tomography lab (now Qurit lab). I would like to thank Kevin Leung for being a wonderful friend and for all the interesting discussions we had. I would like to thank Saeed Ashrafinia, Yuanyuan Gao, and Rongkai Yan for all the good discussions and memories they brought during my first two and half years of my PhD program. I would like to thank Wenbing Lv, Xinchi Hou, Ivan Klyuzhin, Carlos Uribe, Cassandra Miller, Mohammad Salmanpour, Fereshteh Yousefi, Alexander Hart, Roberto Fedrigo, Arash Javanmardi, Adam Watkins, Yas Oloumi, Cariad Knight, Leo Lin, Michael Chang, and Julia Brosch for all the good discussions, interesting activites and happiness they brought during my stay in Vancouver.

I would express my special thanks to my friends who have helped me went through my difficult times and have always brought me much joys in my life. I would like to thank Haochen Kang, Zichao Wang, Xiaoyang Feng, Zijie Wang, Liying Cheng, Muzhi Zhu, Xiaoxing Tan, and Chengbo Zhao for their support during these years.

Last, my deepest gratitude to my parents, Yunce Zhu and Cuiqiao Fu, my sister, Yanqing Zhu, and my girlfriend, Ying Su. This thesis is dedicated to them for their endless love and support.

Contents

Abstract . . . . **ii**

Acknowledgements . . . . **iv**

Contents . . . . **vi**

List of Tables . . . . **x**

List of Figures . . . . **xi**

Chapter 1 Introduction **1**

- 1.1 Motivation and contribution 2
- 1.2 Thesis outline 2

Chapter 2 Background **5**

- 2.1 Fluorescence molecular tomography 5
  - 2.1.1 Overview 5
  - 2.1.2 Photon propagation in biological tissue 8
    - 2.1.2.1 Physical quantities 8
    - 2.1.2.2 The radiative transport equation 10
    - 2.1.2.3 Solving the RTE with deterministic methods 11
    - 2.1.2.4 Solving the RTE with the Monte Carlo method . . . 16
  - 2.1.3 Forward model of FMT 17
  - 2.1.4 Inverse problem of FMT 18

2.2 Positron Emission Tomography 20
2.2.1 PET Physics 20
2.2.2 PET instrumentation and data acquisition 22
2.2.3 PET reconstruction 24
2.2.3.1 Analytical reconstruction 25
2.2.3.2 Statistical reconstruction 26
2.2.4 Image degradation factors 28
2.2.4.1 Attenuation 29
2.2.4.2 Scattered events 30
2.2.4.3 Random events 30
2.2.4.4 Nonuniformity 32
2.2.4.5 Detector dead time 32
2.2.4.6 Inter-crystal blurring 33
2.2.4.7 Positron physics 34
2.2.4.8 Decay correction 35
2.2.4.9 Patient motion 35
2.2.5 Partial volume effect 35

Chapter 3 Incorporating reflection boundary conditions in the Neumann-series-based radiative transport equation 38
3.1 Introduction 39
3.2 Methods 40
3.2.1 The Neumann-series RTE modeling the boundary conditions . 40
3.2.2 Implementation 44
3.2.3 Neumann-series-based reconstruction algorithm for DOI with boundary conditions 46
3.2.4 Simulation experimental setup 49
3.3 Results 51

3.4 Discussions 52

3.5 Conclusions 58

3.A Derivation of the $\boldsymbol{\mathcal{X}\mathcal{R}\mathcal{X}}\Xi$ term 58

3.B Derivation of the $\boldsymbol{\mathcal{X}\mathcal{R}\mathcal{X}\mathcal{K}\mathcal{X}}\Xi$ term 59

3.C Derivation of the $\frac{\partial \boldsymbol{\mathcal{X}}}{\partial \mu_n}\boldsymbol{\mathcal{R}}$ term 61

Chapter 4 Sparsity-initialized maximum-likelihood expectation maximization for fluorescence molecular tomography reconstruction 63

4.1 Introduction 64

4.2 Methods 66

4.2.1 The forward model and reconstruction problem in FMT . . . 66

4.2.2 Modeling Poisson noise in the reconstruction 67

4.2.3 Sparse reconstruction and preconditioning of sensitivity matrix 68

4.2.4 Experiments 69

4.3 Results 74

4.3.1 Uniform cube phantom 74

4.3.2 Digital mouse phantom 76

4.4 Discussion 78

4.5 Conclusion 84

Chapter 5 Partial volume correction for brain PET imaging 85

5.1 Introduction 86

5.2 Methods 91

5.2.1 Regularized Deconvolution Model 91

5.2.2 Parallel Level Sets Regularization 91

5.2.3 Optimization Algorithm 91

5.2.4 Parameter selection 93

5.2.5 Summary of the Algorithm 94
5.3 Experiments 95
5.3.1 Simulation Experiments 95
5.3.2 Application to *In vivo* Human Dataset 99
5.4 Results 101
5.4.1 Results for Simulation Experiments 101
5.4.2 Results for In vivo Human Dataset 103
5.5 Discussion 106
5.6 Conclusion 111

Chapter 6 Conclusions and general discussion 112
6.1 Chapter 3: Incorporating boundary condition for Neumann-series-based RTE 112
6.2 Chapter 4: Reconstruction algorithm for FMT 113
6.3 Chapter 5: Partial volume correction for brain PET imaging 114

Appendix A Application of computational breast phantoms to evaluate reconstruction methods for fluorescence molecular tomography 116
A.1 Introduction 116
A.2 Method 117
A.3 Experiments 118
A.4 Results and conclusions 119

References 120

Curriculum Vitae 154

List of Tables

2-I Properties of commonly used PET radionuclide. 21

3-I RMSE in the output image obtained using the Neumann-series RTE method with and without the reflection boundary conditions 52

4-I Optical properties of digital mouse phantom. 71

4-II Computation time required by MLEM for 1000 iterations. 75

List of Figures

Figure 2-1 Mechanism of fluorescence generation 6

Figure 2-2 An experimental setup for FMT 7

Figure 2-3 Physical interpretation of terms in Neumann-series-based RTE 15

Figure 2-4 Schematic representation of positron annihilation process. . . 21

Figure 2-5 Block detector commonly used in PET systems. 23

Figure 2-6 An example of Radon transform and sinogram 25

Figure 2-7 Illustration of backprojection reconstruction and filtered backprojection. 26

Figure 2-8 Illustration of attenuation in PET. 29

Figure 2-9 Example of scattered event. 30

Figure 2-10 Example of random event. 31

Figure 2-11 Examples of inter-crystal blurring 34

Figure 2-12 Illustration of the causes of partial volume effect 35

Figure 3-1 Physical interpretation of the various operators in the RTE in considered DOI setup. 42

Figure 3-2 A schematic describing the thin boundary layer in the considered DOI setup. 43

Figure 3-3 Physical interpretation of the three reflection terms in the RTE. 44

Figure 3-4 Simulated experimental setup to evaluate the proposed Neumann-series-based RTE model. 49

Figure 3-5 Linear profiles for comparison between MC and the proposed RTE formalism. 51

Figure 3-6 Comparison of the fluence calculated using the MC and proposed Neumann-series RTE-based methods. 53

Figure 3-7 Plot of estimated absorption coefficient and error bars for Neumann-series-based RTE with and without boundary condition. 54

Figure 3-8 The Fresnel reflection coefficient values as a function of the incident angle for various values of refractive index n. 55

Figure 4-1 The experimental setup of cube phantom and its cross section image. 70

Figure 4-2 The experimental setup of digital mouse phantom and its cross section image. 71

Figure 4-3 Cross sections for the cube phantom reconstructed by MLEM with different iteration number with uniform and sparse initial estimates. 73

Figure 4-4 Quantitative results of different reconstruction methods as functions of iteration number for cube phantom. 74

Figure 4-5 Quantitative results of different reconstruction methods as functions of SNR for cube phantom. 75

Figure 4-6 Quantitative results of different reconstruction methods as functions of iteration number for digital mouse phantom. . . . 76

Figure 4-7 Quantitative results of different reconstruction methods as functions of SNR for digital mouse phantom. 77

Figure 4-8 Cross sections of fluorescence target reconstructed with pure sparse reconstruction method and the proposed method with different truncation number K for digital mouse phantom. . . 79

Figure 4-9 Plot of RMSE vs. truncation number for pure sparse reconstruction method and the proposed reconstruction method for different noise levels. 80

Figure 4-10 An example to illustrate the reason that sparsity-initialized MLEM is able to suppress the noise in the background region. 81

Figure 5-1 Transaxial, coronal and sagittal slices of simulated PET phantom, MR phantom and segmented tissue map. 96

Figure 5-2 Transaxial slice of difference image between perfect registered MR image and misregistered MR image. 98

Figure 5-3 Transaxial slice of tissue maps with perfect segmentation and segmentation with different levels of mismatch. 99

Figure 5-4 Transaxial, coronal and sagittal slice of ground truth, uncorrected image and different PVC methods using perfect MR guidance. 101

Figure 5-5 Image patches for different MR-guided PVC methods with different level of registration and segmentation mismatch. . . . 102

Figure 5-6 Bias-noise (cov) trade-off curves for different PVC methods in different ROIs using MR guidance with different levels of registration error. 104

Figure 5-7 PVC results on ^{15}O-H$_2$O scans for an amyloid negative and amyloid positive participant. 105

Figure 5-8 PVC results on ^{11}C-PIB scans for an amyloid negative and amyloid positive participant. 107

Figure 5-9 Mean concentration in different ROIs for different PVC methods for amyloid negative, amyloid positive participants and their difference. 107

Figure 5-10 Normalized histograms of two-sided t-values for the age term in the linear regression model with voxelwise SUVR as the outcome and spatial distribution of the statistically significant voxels for different PVC methods. 110

Figure A-1 Comparison of reconstructed breast phantom with Tikhonov regularization method and the sparse reconstruction method. . 118

Chapter 1

Introduction

Molecular imaging uses specific molecular probes to provide image contrast for non-invasive and real-time visualization of biochemical events at cellular and molecular level within living cells, tissues, and intact subjects [1]. Although it traces its root in nuclear medicine [2–4], the discipline of molecular imaging has been extended to many other imaging modalities such as computed tomography (CT), magnetic resonance imaging, ultrasound imaging and optical imaging [2, 5–8]. Today, molecular imaging has been widely used in both clinical and pre-clinical applications.

The ability to provide quantitative information allows MI to move beyond simply visualizing biochemical process. Quantitative information can provide additional value to MI in applications such as disease diagnosis, clinical decision making and therapy response assessment [9]. Although not all MI methods have traditionally involved quantitative imaging, progresses in hardware and image generation algorithms has enabled quantitative imaging in a wider array of MI disciplines.

Development of quantitative imaging in MI has benefited a lot from advanced image generation algorithms. For example, in nuclear medicine, developments on statistical iterative algorithms and various degradation compensation techniques enable improved quantitative accuracy for emission tomography. As another example, in optical molecular imaging, advanced photon propagation models and reconstruction methods generate images with higher resolution as well as more accurate quantification.

1.1 Motivation and contribution

This dissertation focuses on two specific MI modalities, i.e. fluorescence molecular tomography (FMT) and positron emission tomography (PET). It aims to improve quantitative performance of these two imaging modalities with advanced image generation algorithms. Contributions of this dissertation include:

1) development of a novel Neumann-series-based radiative transport equation that incorporates reflection boundary condition to model photon propagation in optical imaging.

2) Proposal of a reconstruction algorithm for FMT that is based on sparsity-initialized maximum likelihood expectation maximization (MLEM).

3) Development of a magnetic resonance imaging (MRI)-guided partial volume correction method for brain PET imaging that aims to use high-resolution anatomical information to recover quantitative loss caused by limited spatial resolution of PET system.

1.2 Thesis outline

The thesis is organized as follows:

Chapter 2: background

This chapter provides background knowledge to the following three chapters. This chapter has two main sections that discuss detail on FMT and PET separately. In the FMT subsection, we first briefly introduce the fluorescence mechanism and FMT instrumentation. Then we derive photon propagation model based on radiative transport equation (RTE). Diffusion approximation and Monte Carlo methods, as the most commonly used deterministic and stochastic methods, respectively, are introduced. Another deterministic model based on Neumann-series is also introduced. This model is closely related to the content in Chapter 3. Next, we discuss the forward model and inverse problem of FMT. This part is related to the content in Chapter 4. In the PET subsection, we elaborate on PET physics, instrumentation and data acquisition as well as commonly used reconstruction methods, including both analytical and statistical methods. We then discuss several image degradation factors in PET. In the last part of this chapter, we introduce the partial volume effect. The correction method will be discussed in detail in Chapter 5.

Chapter 3: Incorporating reflection boundary conditions in the Neumann-series-based radiative transport equation

In this chapter, we derive the boundary condition for the Neumann-series-based model, which did not consider photon reflection previously. The new model is validated by comparing with Monte Carlo method. In addition, a diffuse optical imaging (DOI) reconstruction technique based on Neumann-series formalism is also proposed to estimate the optical properties of the tissue. These methods were implemented and evaluated using a simulated 3D diffuse optical imaging setup. This work has been presented as two conference proceeding papers [10, 11] and published as a journal paper on Biomedical Optics Express [12].

Chapter 4: Sparsity-initialized maximum-likelihood expectation maximization for fluorescence molecular tomography reconstruction

In this chapter, we discuss a novel reconstruction method for FMT based on sparsity-initialized maximum-likelihood expectation maximization (MLEM). The method is compared with uniform-initialized MLEM and a sparse reconstruction method with preconditioning. Simulation experiments with a homogeneous cubic phantom and a heterogeneous digital mouse phantom are used to evaluate the performance of these methods. This work has been presented at an SPIE conference and an OSA conference

[13, 14] and has been published as a journal paper in Biomedical Optics Express [15].

Chapter 5: Partial volume correction for brain PET imaging

This chapter discusses detail on the topic of partial volume correction (PVC). We first introduce commonly used PVC methods, including reconstruction-based methods and post-reconstruction methods. We then develop an MRI-guided PVC method for brain PET imaging. The method is based on deconvolution model with parallel level set regularization. A non-smooth optimization algorithm based on split Bregman framework is also developed to solve the optimization problem. The proposed method is evaluated on both simulation data and *in vivo* human data. This work has been presented at multiple international conferences [16–19] and has also been submitted for journal publication [20].

Chapter 6: Conclusion and future work

This chapter concludes the thesis and discusses future work.

Appendix A: Application of computational breast phantoms to evaluate reconstruction methods for fluorescence molecular tomography

In this appendix, we apply our proposed FMT reconstruction method for the task of breast tumor detection. We evaluated our method against conventional Tikhonov regularization method using a realistic digital breast phantom. This work has been presented at 2017 Computational Phantom workshop and has been published as a arxiv preprint [21].

Chapter 2

Background

2.1 Fluorescence molecular tomography

2.1.1 Overview

FMT enables quantitative 3D imaging of fluorophores distribution within biological tissue. Since the advent of optical tomography with fluorescence light in mid 1990s [22, 23], this imaging modality has attracted significant research interest due to potential advantages such as high sensitivity, absence of ionizing irradiation, high temporal resolution and relatively low cost [24]. FMT has been used for both preclinical and clinical applications such as pharmaceutical research [8], brain imaging [25, 26], and tumor detection [27, 28].

FMT makes use of fluorescence light for imaging. Fig. 2-1 (a) illustrate the mechanism of fluorescence generation. When illuminating a fluorescence molecule with light source of proper wavelength, it absorbs excitation photon and transmits to higher vibrational energy state from ground state. The molecule then rapidly relaxes to lower vibrational level without radiation, which is known as internal conversion. In the next transition, the molecule falls back to ground state. The difference energy will be released either through heat or by emitting a fluorescence photon. Due to the energy loss during internal conversion, the energy of emitted fluorescence photon is usually lower than that of excitation photon, resulting in a shift of spectrum towards longer

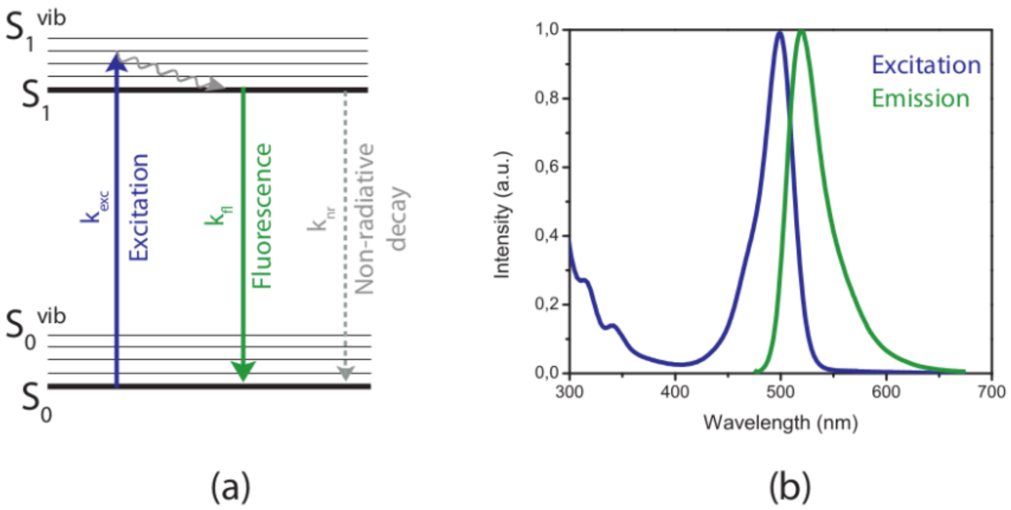


Figure 2-1. (a) Jablonski diagram illustrating fluorescence mechanism. (b) Example of excitation and emission spectra [29].

wavelength, which is known as Stokes shift, as shown in Fig. 2-1 (b). Fluorescence lifetime and quantum yield are two important characteristics of a fluorophore. The former refers to the average time that fluorescence molecule spends in the excited state before returning to the ground state. The later refers to the ratio of the number of emitted photons to the number of absorbed photons [30].

To collect emitted fluorescence light, a typical non-contact experimental setup for small animal FMT is shown in Fig. 2-2. In this setup, the laser source is split into two beams. One is focused to illuminate from the back of the subject to excite fluorescence light. The other is used as front illumination to obtain geometry of subject surface which is used to model photon propagation [32]. The subject is placed on a rotating platform so that signal can be collected from different angle. A CCD camera serves as detector to collect emitted fluorescence light, and a filter is placed in front of the detector to filter out non-fluorescence light. Different setups for FMT have been developed. For example, optic fibers can be used to build a contact system and collect fluorescence signal directly from subject surface [33]. In this type of system, measurement geometry could be transmission-based or reflection-based [26].

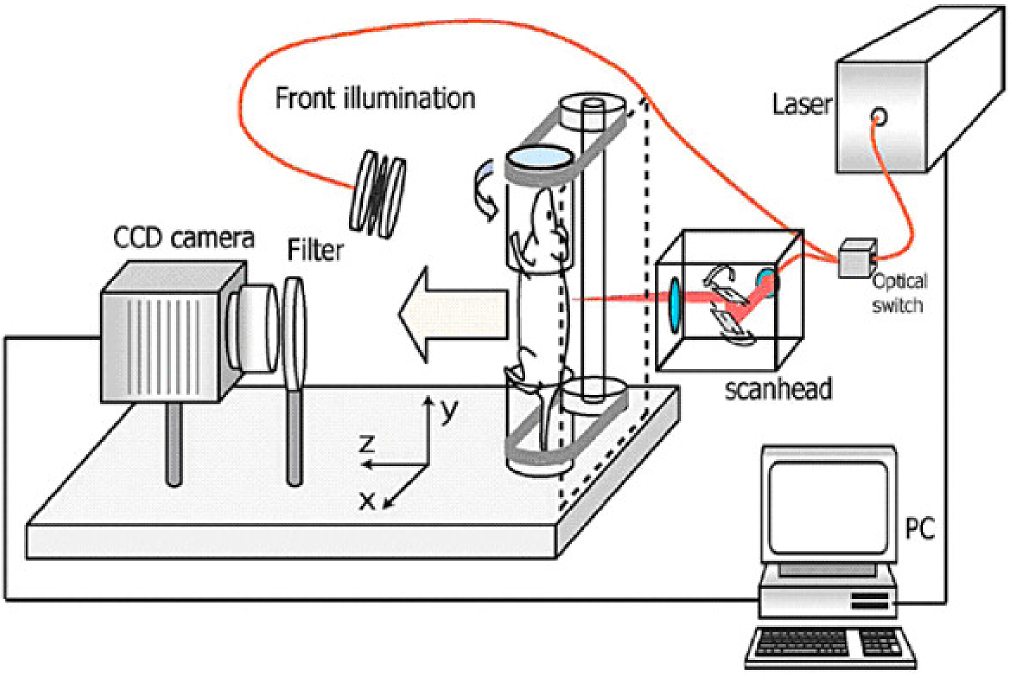


Figure 2-2. An experimental setup for FMT [31].

In addition, subject geometry can be obtained from other modalities. There has been reported FMT-CT and FMT-MRI systems [25, 34, 35], where subject anatomical information is collected through CT or MRI. Anatomical information collected from these systems provide internal structure of subjects and thus heterogeneous tissue map can be applied for more accurate photon propagation modeling. In some other systems, photomultiplier tubes (PMTs) are used as detectors rather than CCD camera [36]. More recently, systems based on wide-field illumination have attracted much research interest, where the focused point laser source is replaced with an expanded wide-field source [37, 38]. This type of systems allows faster data acquisition as well as flexible spatial modulation of illumination pattern which provides improved image quality. Based on the collected data type, these FMT systems can be categorized into three major modes, i.e. continuous wave (CW) mode, frequency-domain (FD) mode and time-domain (TD) mode [39]. CW mode uses laser source with steady illumination power [40]. It is the most commonly used mode due to its low cost and simplicity. Since it only collects intensity data, it can be used to image fluorescence concentration or quantum yield, but not for fluorescence lifetime imaging. In contrast, FD and TD mode allow both concentration imaging and lifetime imaging. In FD mode, frequency modulated (~MHz) source is applied to collect both intensity and phase information [36, 41]. In TD mode, ultra-short (~fs-ps) laser pulse is used as source together with time-resolved detection to collect temporal point spread function information of arrival photon [42, 43].

Photon migration in turbid medium such as biological tissue is complicated by scattering and absorption. In order to reconstruct fluorescence distribution with collected diffuse photons, a forward model is required to describe photon propagation. In the next section, we will discuss details on mathematical modeling that describes photon propagation in biological tissue.

2.1.2 Photon propagation in biological tissue

2.1.2.1 Physical quantities

We start by defining physical quantities that will be used in this section. Let Q denote radiant energy, which is the total energy emitted from a source. If we use V, Ω, E to denote volume, solid angle and energy, respectively, the fundamental quantity used throughout this section is photon distribution function $w(\boldsymbol{r}, \hat{\boldsymbol{s}}, E, t)$, defined as

$$w(\boldsymbol{r}, \hat{\boldsymbol{s}}, E, t) = \frac{1}{E}\frac{\partial^3 Q}{\partial V \partial \Omega \partial E}. \tag{2.1}$$

$w(\boldsymbol{r}, \hat{\boldsymbol{s}}, E, t)\Delta V \Delta\Omega \Delta E$ can be interpreted as the number of photons contained in volume ΔV centered at position $\boldsymbol{r}$, traveling in solid angle $\Delta\Omega$ about direction $\hat{\boldsymbol{s}}$, with energies between E and $E + \Delta E$ at time t. When energy is expressed as eV, the distribution function can be easily related to a more commonly used quantity radiance L with $L = c_m w$, where c_m is the speed of light in the medium.

There are several other important quantities. We define $\Phi = \frac{\partial Q}{\partial t}$ as radient power. Source distribution $\Xi(\boldsymbol{r}, \hat{\boldsymbol{s}}, E, t)$ describes photons injected into a medium and is defined as

$$\Xi(\boldsymbol{r}, \hat{\boldsymbol{s}}, E, t) = \frac{1}{E}\frac{\partial^3 \Phi}{\partial V \partial\Omega \partial E}. \tag{2.2}$$

Irradiance I describes radiant power received by a surface per unit area, defined as

$$I = \frac{\partial \Phi}{\partial A}. \tag{2.3}$$

A similar quantity is fluence rate ϕ, which is also defined as radiant power per unit area, but does not specify the direction. We also define a vector quantity photon flux as

$$\boldsymbol{J}(\boldsymbol{r}, \hat{\boldsymbol{s}}, E, t) = c_m \int_{4\pi} d\Omega \hat{\boldsymbol{s}} w(\boldsymbol{r}, \hat{\boldsymbol{s}}, E, t) \tag{2.4}$$

Scattering cross section σ_{sc} describes the probability of scattering happening for a particle. When a beam of radiation incident on the entity yields a certain scattered flux, the scattering cross section is a measure of scattered flux per unit irradiance. For elastic scattering, which will be considered in this thesis, we have:

$$\sigma_{sc} = \frac{\Phi}{I_0}, \tag{2.5}$$

where I_0 represents normal irradiance on surface. Similarly, we can define absorption cross section σ_{abs} as rate of energy absorption per unit irradiance

$$\sigma_{abs} = \frac{\partial Q_{abs}/\partial t}{I_0}, \tag{2.6}$$

With cross sections defined, we can define absorption coefficient μ_{abs} and scattering coefficient μ_{sc} as

$$\mu_{abs} = n_{abs}\sigma_{abs}, \tag{2.7}$$

$$\mu_{sc} = n_{sc}\sigma_{sc}, \tag{2.8}$$

where n_{abs} and n_{sc} denotes the number of absorbing and scattering particles per unit volume, respectively.

2.1.2.2 The radiative transport equation

Photon propagation within medium is described with the ratiative transport equation (RTE). For optical imaging using diffuse photons such as FMT, it is reasonable to assume mono-energy light source as well as elastic scattering. Thus, we can drop the dependence on energy. With photon distribution function $w(\boldsymbol{r}, \hat{\boldsymbol{s}}, t)$, the RTE is written as

$$\frac{dw}{dt} = \left[\frac{\partial w}{\partial t}\right]_{abs} + \left[\frac{\partial w}{\partial t}\right]_{em} + \left[\frac{\partial w}{\partial t}\right]_{sc} + \left[\frac{\partial w}{\partial t}\right]_{prop}, \tag{2.9}$$

where we have written $w(\boldsymbol{r}, \hat{\boldsymbol{s}}, t)$ as w for short, the right four terms describe photon absorption, emission, scattering and propagation, respectively.

Time-derivative of photon distribution function contributed by absorption is given by

$$\left[\frac{\partial w}{\partial t}\right]_{abs} = -c_m(\boldsymbol{r})\mu_{abs}(\boldsymbol{r})w(\boldsymbol{r}, \hat{\boldsymbol{s}}, t), \tag{2.10}$$

Photon emission is directly given by source distribution $\Xi(\boldsymbol{r}, \hat{\boldsymbol{s}}, t)$, so we have

$$\left[\frac{\partial w}{\partial t}\right]_{em} = \Xi(\boldsymbol{r}, \hat{\boldsymbol{s}}, t), \tag{2.11}$$

The contribution of scattering has two parts. The first part describes photons that scatter away from current direction $\hat{\boldsymbol{s}}$ at position $\boldsymbol{r}$. This part can be written similar as photon absorption:

$$\left[\frac{\partial w}{\partial t}\right]_{out} = -c_m(\boldsymbol{r})\mu_{sc}(\boldsymbol{r})w(\boldsymbol{r}, \hat{\boldsymbol{s}}, t). \tag{2.12}$$

The second part describes photons that scatter from other direction into current direction at position $\boldsymbol{r}$, and is given by

$$\left[\frac{\partial}{\partial t}w(\boldsymbol{r}, \hat{\boldsymbol{s}}, t)\right]_{in} = c_m(\boldsymbol{r})\mu_{sc}(\boldsymbol{r}) \int_{4\pi} d\Omega' p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')w(\boldsymbol{r}, \hat{\boldsymbol{s}}', t), \tag{2.13}$$

where $\hat{\boldsymbol{s}}'$ denotes the direction of incoming photons. $p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}')$ is the scattering phase function, which is commonly chosen as Henyey–Greenstein function [44]

$$p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}') = \frac{1}{4\pi}\frac{1-g^2}{[1+g^2-2g(\hat{\boldsymbol{s}}\cdot\hat{\boldsymbol{s}}')]^{3/2}}, \tag{2.14}$$

where g is anisotropy factor.

The contribution of scattering thus can be written as

$$\begin{aligned}\left[\frac{\partial w}{\partial t}\right]_{sc} = \left[\frac{\partial w}{\partial t}\right]_{out} + \left[\frac{\partial w}{\partial t}\right]_{in} \\ = -c_m(\boldsymbol{r})\mu_{sc}(\boldsymbol{r})w(\boldsymbol{r},\hat{\boldsymbol{s}},t) + c_m(\boldsymbol{r})\mu_{sc}(\boldsymbol{r})\int_{4\pi} d\Omega' p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')w(\boldsymbol{r},\hat{\boldsymbol{s}}',t),\end{aligned} \tag{2.15}$$

Finally, the propagation term is given by

$$\left[\frac{\partial w}{\partial t}\right]_{prop} = -c_m(\boldsymbol{r})\hat{\boldsymbol{s}}\cdot\nabla w(\boldsymbol{r},\hat{\boldsymbol{s}},t). \tag{2.16}$$

We replace the terms in equation (2.9) with equation (2.10), (2.11), (2.15) and (2.16). This gives [45]

$$\begin{aligned}\frac{d}{dt}w(\boldsymbol{r},\hat{\boldsymbol{s}},t) = -c_m(\boldsymbol{r})\mu_t(\boldsymbol{r})w(\boldsymbol{r},\hat{\boldsymbol{s}},t) + \int_{4\pi} d\Omega' K(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}'|\boldsymbol{r})w(\boldsymbol{r},\hat{\boldsymbol{s}}',t) \\ - c_m(\boldsymbol{r})\hat{\boldsymbol{s}}\cdot\nabla w(\boldsymbol{r},\hat{\boldsymbol{s}},t) + \Xi(\boldsymbol{r},\hat{\boldsymbol{s}},t),\end{aligned} \tag{2.17}$$

where $\mu_t(\boldsymbol{r}) = \mu_{sc}(\boldsymbol{r}) + \mu_{abs}(\boldsymbol{r})$ is the total attenuation coefficient, $K(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}'|\boldsymbol{r}) = c_m(\boldsymbol{r})\mu_{sc}(\boldsymbol{r})p(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}')$ is the scattering kernel.

2.1.2.3 Solving the RTE with deterministic methods

Solving the RTE analytically is a difficult problem. Significant researches have been conducted to find approximate solutions for the RTE. A well-established methods to treat the RTE is to expand it with spherical harmonics [46]. Spherical harmonics allow expanding a function based on its angular dependence. If we apply spherical harmonics expansion on the distribution function $w(\boldsymbol{r},\hat{\boldsymbol{s}},t)$, we get

$$w(\boldsymbol{r},\hat{\boldsymbol{s}},t) = \sum_{l=0}^{\infty}\sum_{m=-l}^{m=l} W_{lm}(\boldsymbol{r},t)Y_{lm}(\hat{\boldsymbol{s}}), \tag{2.18}$$

wherer $Y_{lm}(\hat{\boldsymbol{s}})$ is the basis function for spherical harmonics. The coefficients $W_{lm}(\boldsymbol{r},t)$ can be determined by

$$W_{lm}(\boldsymbol{r},t) = \int_{4\pi} d\Omega Y^*_{lm}(\hat{\boldsymbol{s}})w(\boldsymbol{r},\hat{\boldsymbol{s}},t), \tag{2.19}$$

where $Y^*_{lm}(\hat{\boldsymbol{s}})$ is the complex conjugation of $Y_{lm}(\hat{\boldsymbol{s}})$.

Diffuse approximation

The most widely used deterministic model for photon migration modeling is diffuse approximation (DA). If we truncate equation (2.18) at order $l = 1$, we can relate the distribution function with the fluence rate $\phi(\boldsymbol{r}, t)$ and photon flux $\boldsymbol{J}(\boldsymbol{r}, t)$ with

$$w(\boldsymbol{r}, \hat{\boldsymbol{s}}, t) = \frac{1}{4\pi c_m(\boldsymbol{r})} \left[\phi(\boldsymbol{r}, t) + 3\hat{\boldsymbol{s}} \cdot \boldsymbol{J}(\boldsymbol{r}, t)\right]. \tag{2.20}$$

Assuming the source $\Xi(\boldsymbol{r}, \hat{\boldsymbol{s}}, t)$ is isotropic, we drop the angular dependence and get

$$\xi(\boldsymbol{r}, t) = 4\pi\Xi(\boldsymbol{r}, \hat{\boldsymbol{s}}, t). \tag{2.21}$$

Substituting equation (2.20), (2.21) into (2.17) and integral over the 4π space, we obtain

$$\frac{1}{c_m(\boldsymbol{r})}\frac{\partial\phi(\boldsymbol{r}, t)}{\partial t} + \mu_{abs}(\boldsymbol{r})\phi(\boldsymbol{r}, t) + \nabla \cdot \boldsymbol{J}(\boldsymbol{r}, t) = \xi(\boldsymbol{r}, t) \tag{2.22}$$

In strongly scattering medium, we can apply *Fick's law*, which says the photon flux can be approximated with the gradient of fluence rate as

$$\boldsymbol{J}(\boldsymbol{r}, t) = -D(\boldsymbol{r})\nabla\phi(\boldsymbol{r}, t), \tag{2.23}$$

where $D = \frac{1}{3(\mu_{abs}+\mu'_{sc})}$ is diffusion coefficient, $\mu'_{sc} = (1 - g)\mu_{sc}$ is reduced scattering coefficient. If we use *Fick's law* in equation (2.22), we obtain the well-known diffusion equation

$$\frac{1}{c_m(\boldsymbol{r})}\frac{\partial\phi(\boldsymbol{r}, t)}{\partial t} + \mu_{abs}(\boldsymbol{r})\phi(\boldsymbol{r}, t) - \nabla \cdot [D(\boldsymbol{r})\nabla\phi(\boldsymbol{r}, t)] = \xi(\boldsymbol{r}, t) \tag{2.24}$$

Equation (2.24) can be solved with numerical method such as finite element method to obtain the optical field in complicated geometry [47, 48]. Notice we have made several assumptions in diffuse approximation. As a result, the diffuse approximation is valid when $\mu'_{sc} \gg \mu_{abs}$ and the observation point is far from the boundary and the source [49].

Neumann-series method

As discussed in the previous section, there are several limitations to the DA method due to assumptions made in its derivation. As a result, DA-based methods are not accurate, for example, in media with low scattering but high absorption or small geometry [50–53]. To overcome these issues, we now discuss a Neumann-series method that aims to solve the RTE. Neumann-series method solves the RTE with integral form. Unlike methods using differential form of the RTE that require solving multiple coupled equation, Neumann-series method only require solving a single equation [54]. In addition, this method has a clear physical interpretation for the RTE solution, as we will show later. We next describe how to derive Neumann-series method to solve the RTE.

We first apply Fourier transform on equation (2.17). This gives

$$\begin{aligned}\hat{\boldsymbol{s}}\cdot\nabla w(\boldsymbol{r},\hat{\boldsymbol{s}},\nu)+\left(\mu_t(\boldsymbol{r})+\frac{i\nu}{c_m(\boldsymbol{r})}\right)w(\boldsymbol{r},\hat{\boldsymbol{s}},\nu)=\frac{1}{c_m(\boldsymbol{r})}\int_{4\pi}d\Omega' K(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}'|\boldsymbol{r})w(\boldsymbol{r},\hat{\boldsymbol{s}}',\nu)\\+\frac{1}{c_m(\boldsymbol{r})}\Xi(\boldsymbol{r},\hat{\boldsymbol{s}},\nu),\end{aligned}\tag{2.25}$$

wherer ν is frequency. For steady-state source, $\nu = 0$. The following discussion will be based on the assumption of steady-state source. The model can be easily extended to general frequency domain by setting $\mu_t'(\boldsymbol{r}) = \mu_t(\boldsymbol{r}) + \frac{i\nu}{c_m(\boldsymbol{r})}$. With steady-state source, we have

$$\hat{\boldsymbol{s}}\cdot\nabla w(\boldsymbol{r},\hat{\boldsymbol{s}})+\mu_t(\boldsymbol{r})w(\boldsymbol{r},\hat{\boldsymbol{s}})=\frac{1}{c_m(\boldsymbol{r})}\int_{4\pi}d\Omega' K(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}'|\boldsymbol{r})w(\boldsymbol{r},\hat{\boldsymbol{s}}')+\frac{1}{c_m(\boldsymbol{r})}\Xi(\boldsymbol{r},\hat{\boldsymbol{s}}).\tag{2.26}$$

If we perform integration along the path in $\hat{\boldsymbol{s}}$ direction, we obtain [45]

$$\begin{aligned}w(\boldsymbol{r},\hat{\boldsymbol{s}})=\frac{1}{c_m(\boldsymbol{r})}\int_0^\infty dl\,\Xi(\boldsymbol{r}-\hat{\boldsymbol{s}}l,\hat{\boldsymbol{s}})\exp\left[-\int_0^l dl'\mu_t(\boldsymbol{r}-\hat{\boldsymbol{s}}l')\right]\\+\frac{1}{c_m(\boldsymbol{r})}\int_0^\infty dl[\boldsymbol{\mathcal{K}}w](\boldsymbol{r}-\hat{\boldsymbol{s}}l,\hat{\boldsymbol{s}})\exp\left[-\int_0^l dl'\mu_t(\boldsymbol{r}-\hat{\boldsymbol{s}}l')\right],\end{aligned}\tag{2.27}$$

where $\boldsymbol{\mathcal{K}}$ is scattering operator and is defined as

$$[\boldsymbol{\mathcal{K}}w](\boldsymbol{r},\hat{\boldsymbol{s}})=\int_{4\pi}d\Omega' K(\hat{\boldsymbol{s}},\hat{\boldsymbol{s}}'|\boldsymbol{r})w(\boldsymbol{r},\hat{\boldsymbol{s}}').\tag{2.28}$$

We define another operator $\boldsymbol{\mathcal{X}}$ from equation (2.27) as

$$[\boldsymbol{\mathcal{X}}w](\boldsymbol{r},\hat{\boldsymbol{s}}) = \frac{1}{c_m(\boldsymbol{r})}\int_0^\infty dl w(\boldsymbol{r}-\hat{\boldsymbol{s}}l,\hat{\boldsymbol{s}})\exp\left[-\int_0^l dl'\mu_t(\boldsymbol{r}-\hat{\boldsymbol{s}}l')\right]. \tag{2.29}$$

The operator $\boldsymbol{\mathcal{X}}$ defines an integral transform which is known as x-ray transform [45]. This operator describes the attenuation of photons when propagating along $\hat{\boldsymbol{s}}$ direction due to absorption and scattering away from current direction. We refer to this operator as attenuation operator. Use equation (2.28) and (2.29), we can write equation (2.27) in operator form as

$$w = \boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}w. \tag{2.30}$$

An equivalent form is

$$(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}})w = \boldsymbol{\mathcal{X}}\Xi. \tag{2.31}$$

Expand the above equation with Neumann-series and we obtain

$$\begin{aligned} w &= (\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}})^{-1}\boldsymbol{\mathcal{X}}\Xi \\ &= \boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi + \ldots \end{aligned} \tag{2.32}$$

Each term in this solution has an intuitive interpretation. The first term $\boldsymbol{\mathcal{X}}\Xi$ describes the contribution of ballistic photons that propagates through a medium without scattering. Terms with operator $[\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}]^n$ describe contribution from photons that scatter n times [54]. This is illustrated in Fig. 2-3.

With spherical harmonics defined in (2.18) and (2.19), we transform operator $\boldsymbol{\mathcal{K}}$ and $\boldsymbol{\mathcal{X}}$ in spherical harmonics basis. We first define an operator $\boldsymbol{\mathcal{Y}}$ that maps the distribution function $w(\boldsymbol{r},\hat{\boldsymbol{s}})$ to its spherical harmonics coefficient $W_{lm}(\boldsymbol{r})$. This is written as

$$[\boldsymbol{\mathcal{Y}}w]_{lm}(\boldsymbol{r}) = \int_{4\pi} d\Omega Y_{lm}^*(\hat{\boldsymbol{s}})w(\boldsymbol{r},\hat{\boldsymbol{s}},t) = W_{lm}(\boldsymbol{r}). \tag{2.33}$$

We can also define its inverse operator as

$$[\boldsymbol{\mathcal{Y}}^{-1}W](\boldsymbol{r},\hat{\boldsymbol{s}}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{m=l} W_{lm}(\boldsymbol{r},t)Y_{lm}(\hat{\boldsymbol{s}}) = w(\boldsymbol{r},\hat{\boldsymbol{s}}). \tag{2.34}$$

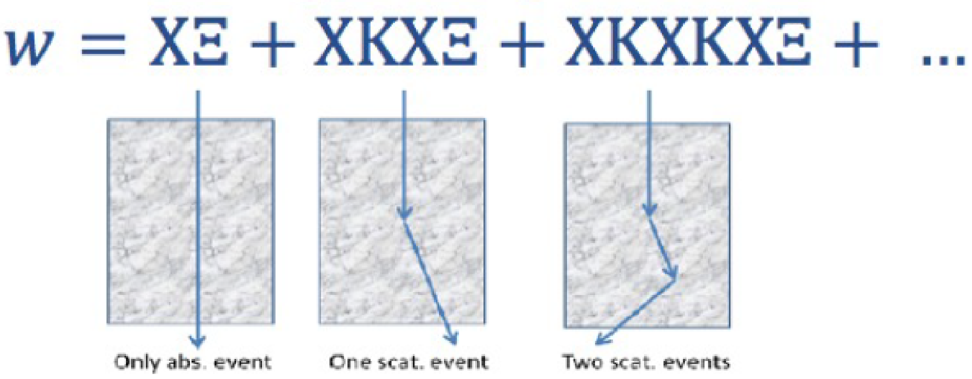


Figure
the Neumann series RTE [11].

Applying $\boldsymbol{\mathcal{Y}}$ on both side of (2.30), we get

$$\boldsymbol{\mathcal{Y}}w = \boldsymbol{\mathcal{Y}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{Y}}^{-1}\xi + \boldsymbol{\mathcal{Y}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{Y}}^{-1}\boldsymbol{\mathcal{Y}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{Y}}^{-1}W, \tag{2.35}$$

where $\xi = \boldsymbol{\mathcal{Y}}\Xi$ is spherical harmonics coefficients of the source. We now define the transformed operator of $\boldsymbol{\mathcal{K}}$ and $\boldsymbol{\mathcal{X}}$ as $\boldsymbol{\mathcal{D}} = \boldsymbol{\mathcal{Y}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{Y}}^{-1}$ and $\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{Y}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{Y}}^{-1}$. It can be shown that the transformed scattering operator $\boldsymbol{\mathcal{D}}$ is

$$[\boldsymbol{\mathcal{D}}W]_{lm}(\boldsymbol{r}) = c_m(\boldsymbol{r})\mu_{sc}(\boldsymbol{r}) \sum_{l',m'} g^l \delta_{ll'}\delta_{mm'}W_{l'm'}(\boldsymbol{r}), \tag{2.36}$$

where δ_{mn} denote the Kronecker delta function. The transformed attenuation operator $\boldsymbol{\mathcal{A}}$ is

$$\begin{aligned}[\boldsymbol{\mathcal{A}}W]_{lm}(\boldsymbol{r}) = &\frac{1}{c_m(\boldsymbol{r})} \sum_{l',m'} \int_{\infty} d\boldsymbol{r}' \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|^2} Y^*_{lm}\left(\frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|}\right) \\ &\times Y_{lm}\left(\frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|}\right) \exp\left[-\int_0^{|\boldsymbol{r}-\boldsymbol{r}'|} dl' \mu_t(\boldsymbol{r} - \frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|} l')\right],\end{aligned} \tag{2.37}$$

We can then write the Neumann-series-based RTE in spherical harmonics as

$$W = \boldsymbol{\mathcal{A}}\xi + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{A}}\xi + \boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{A}}\boldsymbol{\mathcal{D}}\boldsymbol{\mathcal{A}}\xi + \dots \tag{2.38}$$

Based on equation (2.38), Neumann-series-based RTE has been successfully implemented in both uniform medium [54] and non-uniform medium [55].

Other analytical method

Many other analytical methods have also been developed in order to overcome the drawbacks of DA. Instead of using the first order spherical harmonics, high order methods have been developed [56–59]. This type of methods is known as P_N method. Methods using simplified harmonics (SP_N) were also developed in order to address the issue of mathematical complexity and computationally demanding in P_N method [52, 60]. Apart from spherical harmonics expansion, other methods have been developed to solve the RTE. Klose et. al. [61] use discrete ordinate method to discretize angular space. Pulkkinen and Tarvainen [62] use truncated Fourier series to expand RTE.

2.1.2.4 Solving the RTE with the Monte Carlo method

Monte Carlo (MC) method is another commonly used method to solve RTE. MC method simulates photon migration stochastically. In general, MC methods apply random sampling and statistical modeling to estimate mathematical functions. In a typical MC simulation, a system is modeled as probability density functions (PDFs). Repeated samples are obtained from these PDFs. Finally, statistics of interest will be computed based on these samples [63]. Specifically, for photon migration problem, different parameters which characterize photon behavior, such as propagation step size and scattering angle, obey certain probability distributions that are defined based on optical properties of different medium [64]. After sufficient number of photons sampled, the cumulative distribution of all photon paths provide accurate solution to RTE.

Although MC method is accurate, it requires large number of photons to be simulated in order to obtain a stable solution, which made it very slow in early years. More recently, different strategies have been developed to accelerate MC simulation so that it can be applied to obtain the forward model of optical imaging with diffuse light. Strategies such as scaling [65] and perturbation [66] try to maximize the use of information collected in certain MC simulations to deduce photon distribution in medium with different optical properties. Other strategies such as filtering [67] try to improve signal-to-noise ratio (SNR) and reduce the number of photons required for a single MC simulation. The emergence of graphics processing units (GPUs) and parallel computing techniques have also contributed to significantly speed up MC simulation [68].

2.1.3 Forward model of FMT

For steady-state source, solving the RTE gives us the Green's function $G(\boldsymbol{r}, \boldsymbol{r}')$, which describes fluence rate at position $\boldsymbol{r}'$ due to point source at $\boldsymbol{r}$. Photon propagation in FMT is a two-stage process. In the first stage, excitation photons propagate from source into tissue. In the second stage, emitted fluorescence photons propagate from fluorophores to detectors outside the boundary of the subject. We use $G_{ex}(\boldsymbol{r}, \boldsymbol{r}')$ and $G_{em}(\boldsymbol{r}, \boldsymbol{r}')$ to denote Green's function of excitation photon and emission photon, respectively. Then the fluence rate for excitation photon at position $\boldsymbol{r}$ is

$$\phi_{ex}(\boldsymbol{r}) = \int_{\Lambda} G_{ex}(\boldsymbol{r}_s, \boldsymbol{r}) q(\boldsymbol{r}_s) d\boldsymbol{r}_s, \tag{2.39}$$

where Λ is the domain of the subject. $q(\boldsymbol{r}_s)$ is excitation source, and $\boldsymbol{r}_s$ is source position. For emitted fluorescence, we have [69]

$$\phi_{em}(\boldsymbol{r}_d) = \int_{\Lambda} G_{em}(\boldsymbol{r}, \boldsymbol{r}_d) \eta \mu_{af}(\boldsymbol{r}) \phi_{ex}(\boldsymbol{r}) d\boldsymbol{r}, \tag{2.40}$$

where η is quantum yield of fluorophore and $\mu_{af}(\boldsymbol{r})$ is absorption coefficient of fluorophores at excitation wavelength. If we discretize Λ into N voxels, we obtain the linear matrix equation for the forward model:

$$\mathbf{\Phi} = \mathbf{G}\boldsymbol{x}, \tag{2.41}$$

where

$$\mathbf{G} = \begin{bmatrix} g^{1}_{em,1}\phi^{1}_{em,1} & \cdots & g^{1}_{em,N}\phi^{1}_{em,N} \\ \vdots & & \vdots \\ g^{N_d}_{em,1}\phi^{1}_{em,1} & \cdots & g^{N_d}_{em,N}\phi^{1}_{em,N} \\ g^{1}_{em,1}\phi^{2}_{em,1} & \cdots & g^{1}_{em,N}\phi^{2}_{em,N} \\ \vdots & & \vdots \\ g^{N_d}_{em,1}\phi^{N_s}_{em,1} & \cdots & g^{N_d}_{em,N}\phi^{N_s}_{em,N} \end{bmatrix}$$

is the sensitivity matrix of the system, $\mathbf{\Phi}$ is an $M \times 1$ vector denoting detector measurements. N_s and N_d are number of sources and detectors, respectively, and $M = N_s \times N_d$ is the total number of measurements, $\boldsymbol{x} = \eta \boldsymbol{\mu}_{af}$ is the unknown fluorescence yield.

2.1.4 Inverse problem of FMT

The inverse problem is to reconstruct $\boldsymbol{x}$ from equation (2.41) given detector measurements $\mathbf{\Phi}$ and system matrix $\mathbf{G}$. Since biological tissue is highly scattering, and the number of measurements is usually smaller compared to the number of unknowns. The inverse problem of FMT is highly ill-posed. In order to reduce ill-posedness of the problem, Regularization is often used. The following equation shows a typical least-square regularization model that has been widely used for FMT reconstruction [43, 70–72].

$$\min_{\boldsymbol{x}} \|\mathbf{G}\boldsymbol{x} - \mathbf{\Phi}\|_2^2 + \lambda R(\boldsymbol{x}), \tag{2.42}$$

where $R(\boldsymbol{x})$ is certain regularization function, and λ is regularization parameter.

Various regularization methods have been developed. The most widely used one is Tikhoniv regularization [73]. With Tikhonov regularization, equation (2.42) becomes

$$\min_{\boldsymbol{x}} \|\mathbf{G}\boldsymbol{x} - \mathbf{\Phi}\|_2^2 + \lambda \|\mathbf{\Gamma}\boldsymbol{x}\|_2^2, \tag{2.43}$$

where $\mathbf{\Gamma}$ is a weighting matrix. When no prior information is available, $\mathbf{\Gamma} = \boldsymbol{I}$ is identity matrix. Equation (2.43) becomes l_2 regularization [74]. On the other hand, $\mathbf{\Gamma}$ can be designed for spatial adaptive regularization [71] or to incorporate anatomical prior information [70]. By setting the gradient of objective function in (2.43) to zero, we obtain an explicit solution as

$$\hat{\boldsymbol{x}} = \left(\mathbf{G}^T\mathbf{G} + \lambda \mathbf{\Gamma}^T\mathbf{\Gamma}\right)^{-1} \mathbf{G}^T\mathbf{\Phi}. \tag{2.44}$$

Computing the inverse matrix in (2.44) directly can be slow if matrix size is large. Instead, equation (2.43) can be solved iteratively with numerical method such as gradient descent method [75]. Some methods tries to solve the linear equation (2.41) directly, such as algebraic reconstruction technique (ART), conjugate gradient method and truncated singular value decomposition (TSVD) [31, 46, 76]. In these methods, early stopping is commonly used in iterative methods to obtain meaningful solution. The step size and truncation number in these methods play a similar role as the regularization parameter. Discussion on their relationship with Tikhonov regularization can be found in [77, 78].

In practice, Tikhonov regularization often tends to generate over-smoothed images. Since fluorophores usually concentrate in local regions, sparsity can be exploit to improve image quality. Different regularization functions encouraging sparsity have been studied in recent years, such as l_p $(0 < p \leq 1)$ norm [72, 79, 80], log function [81], total variation [82, 83] and group sparsity functions [84, 85].

In addition, the advent of multimodality systems also encourages the development of FMT reconstruction algorithm. As we discussed earlier in this chapter, multimodality systems such as FMT-CT and FMT-MRI provide structural information, which can be used to provide more accurate photon propagation modeling. On the other hand, such structural information can be incorporated during reconstruction process to improve image quality. Apart from Tikhonov regularization, where anatomical information can be used to design weighting matrix, advanced reconstruction algorithms have been proposed to incorporate structural information, such as anisotropic diffusion regularization [86], group-sparsity function [87], and kernel method [88].

There are other reconstruction methods. For example, instead of treating FMT reconstruction as pixel-based parameter identification problem, shape-based methods try to reconstruct the support of fluorescence distribution [89, 90]. More recently, deep-learning-based reconstruction methods have also been developed which maps boundary measurements directly to inside fluorescence distribution [91, 92].

2.2 Positron Emission Tomography

PET uses positron emitting radionuclides for imaging [93, 94]. As elaborated below, PET imaging involves detection of anti-parallel photon pairs generated following positron annihilation, followed by reconstruction of spatial distributions for radiotracers based on these data. Compared to another widely used emission tomography modality, namely single-photon emission computed tomography (SPECT), that utilize radionuclides with single or multiple uncorrelated gamma photon emissions, the feature of imaging with simultaneous annihilation photons propagating in anti-parallel directions provides several advantage to PET imaging, such as higher sensitivity and easier attenuation correction [95, 96]. Today, PET imaging is actively and clinically performed in a wide range of applications including oncology, neurology, and cardiology [97–99]. In this section we will briefly introduce background knowledge for PET. We start this with more elaborate introduction into the physics of PET imaging.

2.2.1 PET Physics

The process of positron annihilation is illustrated in Fig. 2-4. For radioactive decay with positron emission, a proton in the nucleus is transformed into a neutron and a positively charged electron. For a typical radionuclide ${}_Z^AX$ with A denoting the total number of protons and neutrons, and Z representing the number of protons, the above process is described as

$$ {}_Z^AX \longrightarrow {}_{Z-1}^{A} X + \beta^+ + v, \tag{2.45}$$

where v represents a neutrino. The ejected positron travels a short distance and loses its kinetic energy in collisions with atoms of surrounding materials. The distance for certain radionuclide that positron travels after ejection is known as *positron range*. The positron then interacts with an electron through annihilation reaction, generating two 511 keV γ photons traveling in almost 180° opposing directions and detected by

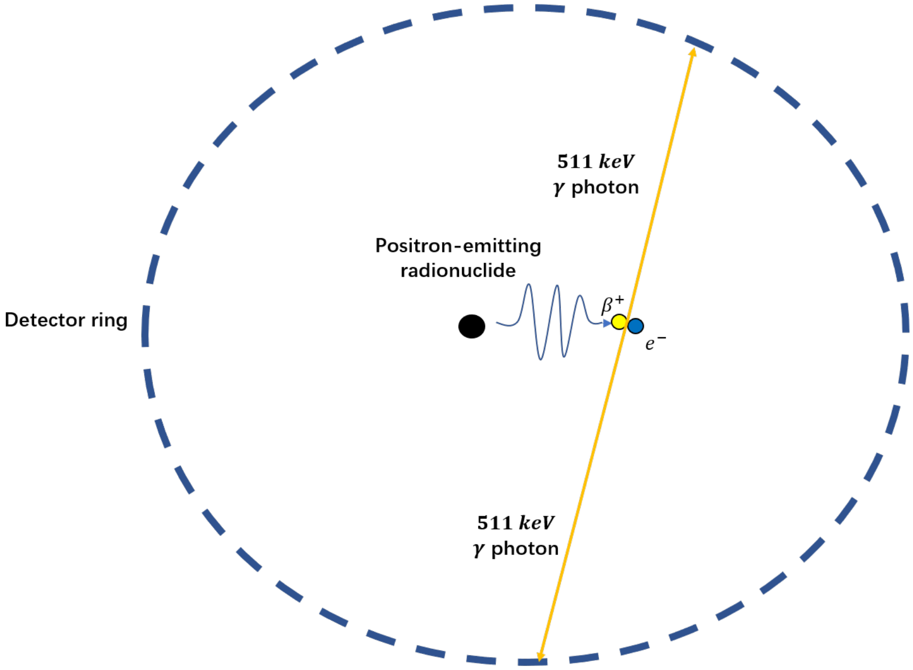

511 keV
γ photon
Positron-emitting
radionuclide
Detector ring
β+
e−
511 keV
γ photon

Figure 2-4. Schematic representation of positron annihilation process.

Table 2-I. Properties of commonly used PET radionuclide [94, 100].

Radionuclide	Half life (mins)	Positron range in water (mm)
^{18}F	110	0.6
^{11}C	20.4	1.1
^{13}N	9.97	1.5
^{15}O	2.03	2.5
^{68}Ga	67.71	2.9
^{82}Rb	1.25	5.9

detector ring in PET scanner.

The radioactive decay of positron emitter follows exponential law and can be described as

$$N_t = N_0 \exp\left(-\lambda t\right), \tag{2.46}$$

where N_0 is the initial number of radioactive atoms, N_t is the number of radioactive atoms at time t, and λ is decay constant. An important parameter for a radionuclide is *half life* $T_{\frac{1}{2}}$, which describes the time required for a radionuclide to decay to 50% of its initial activity level. Its relation with decay constant λ is

$$T_{\frac{1}{2}} = \frac{\ln 2}{\lambda}. \tag{2.47}$$

Table 2-I shows properties of commonly used radionuclide in PET.

2.2.2 PET instrumentation and data acquisition

After arriving at detectors, the 511 keV γ photons interact with detector material, which is made of scintillation crystals. During this process, while most energy of the γ photon will be released as thermal energy, a portion of the energy is released as low-energy photon at wavelength of visible light [94]. In order to effectively stop high-energy gamma ray photons and generate signal carrying accurate information of location of incident photons, deposited energy, as well as timing information, an ideal scintillator should have short decay time, high light output, be dense, and cheap to produce [101]. Commonly used detector materials include bismuth germanate (BGO), lutetium oxyorthosilicate (LSO) and Lutetium-yttrium oxyorthosilicate (LYSO). The low-energy photons produced in scintillation crystals are then channeled into PMTs, which amplify the signal and convert it to electrical signal. Instead of coupling each scintillation detector to a PMT, modern PET systems use block detector, as shown in Fig. 2-5. The use of block detectors allows using smaller scintillation detectors as well as reducing the number of PMTs. This reduces the cost while improving spatial resolution of PET system [102, 103]. Positioning logic circuit is used to determine the coordinates of events on block detector plane.

In order to localize radionuclide with the photon pair generated through positron annihilation, PET system uses a coincidence detection circuit to judge whether two photons arrive at detector “simultaneously”. By setting a time window, coincidence detection circuit will declare an event if the two photons arrive two detectors within this time window. This mechanism is known as annihilation coincidence detection (ACD) [105]. The line linking the two detectors is called line of response (LOR). By

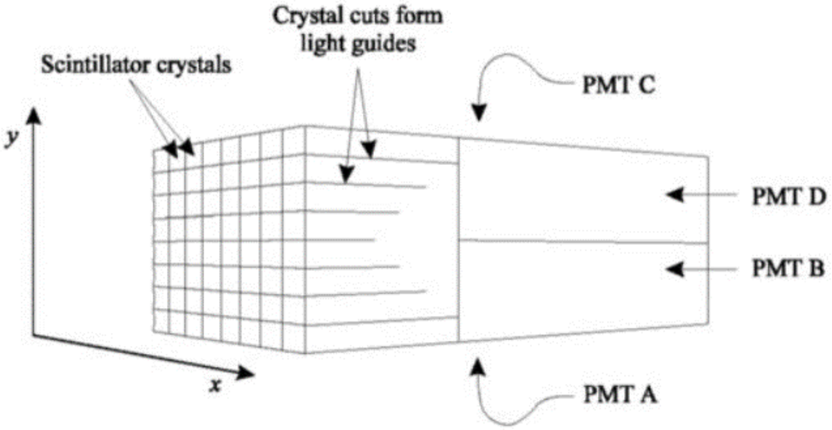


Figure 2-5. Block detector commonly used in PET systems [104].

using ACD, PET system could localize LOR without using a physical collimator. This increases the sensitivity of PET system compared to SPECT system.

Besides being used to judge if an event occurs, the arrival times of photon pair provide more information. If we consider the difference of arrival time, which is known as time-of-flight (TOF) information, theoretically the location of annihilation event could be constrained to a point rather than a line, thus distribution of radiotracer could be obtained without performing reconstruction [106]. In practice, however, current commercial TOF PET scanners usually have timing resolution ranging from about 200-550 ps [107–109]. Due to this limited timing resolution, we currently cannot obtain the exact positron emitter position only based on this information, but the additional TOF information still improves SNR of reconstructed images [110, 111].

Data acquisition in PET can be categorized as 2D mode and 3D mode [112]. For 2D mode, axial collimators are placed between slice of detector rings so that only photons from the same slice or adjacent slices are allowed. In 3D mode, collimators between slices are removed and coincidence events are also recorded along the oblique planes. Compared to 2D mode, 3D acquisition improves sensitivity by a factor of around 5-7 [113, 114], although this benefit is complicated by degradation factors such as increased dead time loss, random coincidence and scattered events [115].

The measurements obtained from detectors are stored either as list-mode data or sinogram data. In list-mode data, detector coordinate of LOR, together with other necessary information, are stored in data packet for each coincidence event [116]. Such data can be directly used for reconstruction [117]. In sinogram data, detector measurements are summed along each LOR and binned into sinogram space based on predefined projection angle, detector coordinate and time frame [118].

Today, commercial PET systems are commonly combined with CT or MR systems. Due to the limited resolution of PET, in many cases it is desirable to localize PET activity with the help of high-resolution anatomic structures from CT or MR [119]. In addition, the anatomical information provided by CT or MR system can also be used to perform attenuation correction [120, 121] or to improve image quality by incorporating such information within the image generation process [112, 122].

2.2.3 PET reconstruction

To describe the process of data acquisition through projection, we first introduce *Radon transform*. The left part of Fig. 2-6 shows an example of projecting a 2D Shepp-Logan phantom to sinogram space. If we use $f(x, y)$ to denote the object, $g(l, \theta)$ to denote the sinogram data, for a line $\{(x, y)|x\cos\theta + y\sin\theta = l\}$, the projection process can be described with the following line integral

$$g(l, \theta) = [\boldsymbol{\mathcal{R}} f](l, \theta) = \int_{\infty}\int_{\infty} f(x, y)\delta(x\cos\theta + y\sin\theta - l)dxdy, \tag{2.48}$$

where x, y are coordinates in object space, θ is projection angle and l is the lateral position. This transform is known as *Radon transform* and we use $\boldsymbol{\mathcal{R}}$ to denote this operation. The right part of Fig. 2-6 shows the sinogram data after Radon transform of the phantom.

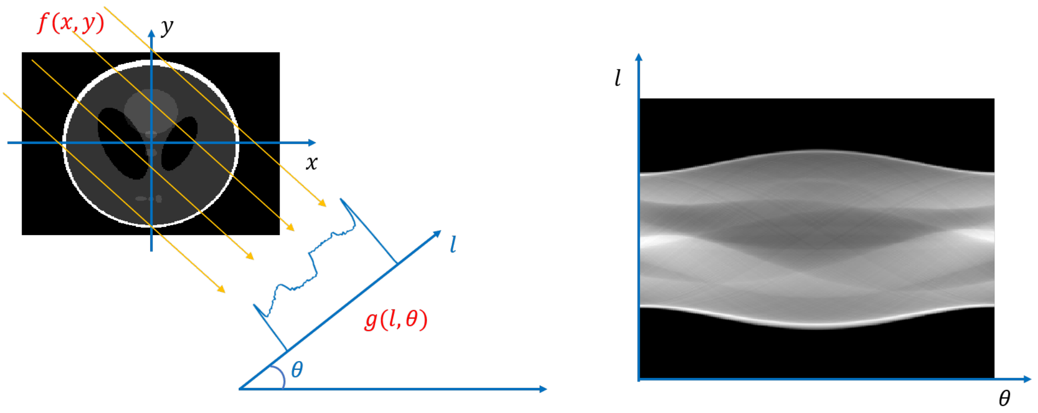


Figure 2-6. An example of Radon transform (left) and sinogram (right).

2.2.3.1 Analytical reconstruction

In order to reconstruct the object, a naive way is to project the sinogram data into object space, and sum them up over all the angles, as illustrated in Fig. 2-7. Mathematically, this can be described as

$$\hat{f}(x,y) = \int_0^{\pi} g(x\cos\theta + y\sin\theta, \theta)d\theta. \tag{2.49}$$

This transform is the adjoint operation of Radon transform, and will be referred to as *backprojection operation*. As shown in Fig. 2-7, reconstruction obtained with backprojection operation produces blurry image.

With central slice theorem, which states the 1D Fourier transform of a projection is a slice of the 2D Fourier transform of the object, we can derive a more commonly used analytical reconstruction method, known as filtered backprojection (FBP). The detail of derivation can be found in [123, 124]. If we use $\boldsymbol{\mathcal{F}}_{1D}$ to denote 1D Fourier transform and $\boldsymbol{\mathcal{R}}^*$ to denote backprojection operator, FBP is written as

$$\begin{aligned}\hat{f}(x,y) &= \boldsymbol{\mathcal{R}}^*\boldsymbol{\mathcal{F}}_{1D}^{-1}[|\rho|\boldsymbol{\mathcal{F}}_{1D}g(l,\theta)] \\ &= \int_0^{\pi}\left[\int_{\infty}|\rho|G(\rho,\theta)\exp\left(2\pi j\rho l\right)d\rho\right]_{l=x\cos\theta+y\sin\theta} d\theta,\end{aligned} \tag{2.50}$$

where $G(\rho,\theta) = \boldsymbol{\mathcal{F}}_{1D}g(l,\theta)$. In equation (2.50), $|\rho|$ plays the role as a high-pass filter and is known as ramp filter. In practice, the ramp filter can be modified to control

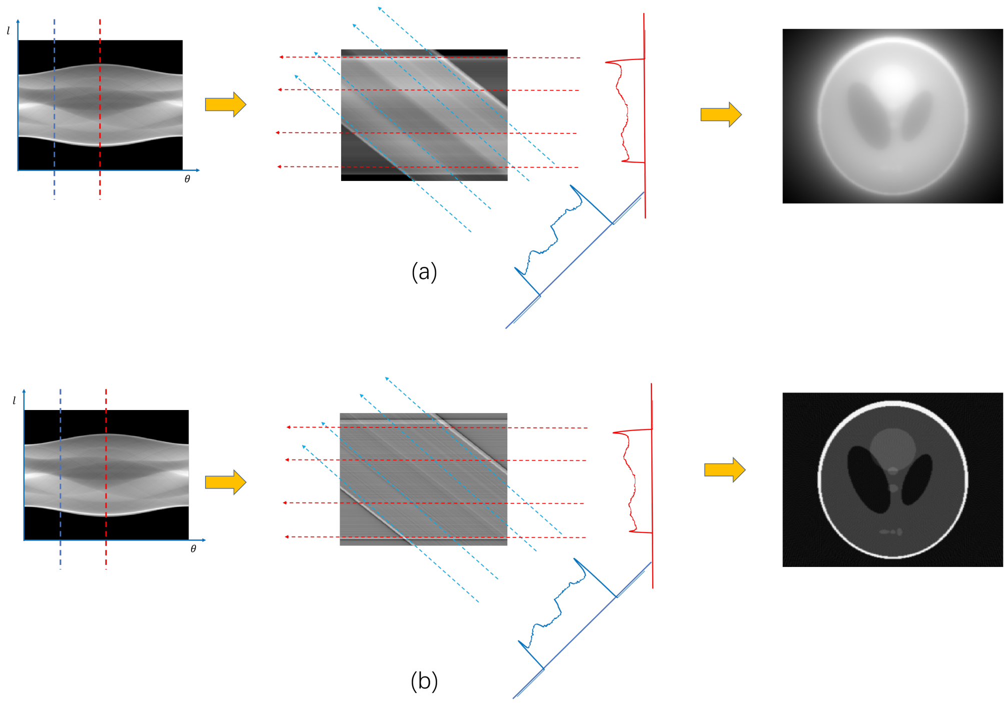


Figure 2-7. Illustration of (a) backprojection reconstruction and (b) filtered backprojection.

noise [112]. Fig. 2-7 (b) gives an example of reconstruction with FBP algorithm. FBP was used for PET reconstruction for many years but was later replaced with statistical iterative methods since they provide more accurate noise model as well as flexibility to incorporate knowledge of PET system and prior information.

2.2.3.2 Statistical reconstruction

The detector measurements obtained in PET system obey Poisson distribution. A statistical iterative algorithm for PET reconstruction, known as *maximum-likelihood expectation maximization* (MLEM), can be derived based on this noise model. To show this, we first perform discretization on detector measurements and image. Let $\boldsymbol{g} = [g_1, g_2, \cdots, g_M]^T$ denote detector measurements, where M is the total number of sinogram bins, $\boldsymbol{f} = [f_1, f_2, \cdots, f_N]^T$ represent PET image, with N representing the number of voxels. Let $\boldsymbol{P} \in \mathbb{R}^{M \times N}$ be the projection matrix. For now we do not consider any degradation factor, so $\boldsymbol{P}$ just performs Radon transform. If we use $p(x|y)$ to represent the probability of variable x conditioned on variable y, the likelihood function of PET image $\boldsymbol{f}$ is

$$p(\boldsymbol{g}|\boldsymbol{f}) = \prod_{m=1}^{M} \exp\left[-(\boldsymbol{P}\boldsymbol{f})_m\right] \frac{(\boldsymbol{P}\boldsymbol{f})_m^{g_m}}{g_m!}, \tag{2.51}$$

If we take the logarithm of the likelihood function, we have

$$\log p(\boldsymbol{g}|\boldsymbol{f}) = \sum_{m=1}^{M} \left\{-(\boldsymbol{P}\boldsymbol{f})_m + g_m \ln\left[(\boldsymbol{P}\boldsymbol{f})_m\right] - \ln g_m!\right\}. \tag{2.52}$$

The first order derivative of the log-likelihood function is given by

$$\frac{\partial}{\partial f_n} \log p(\boldsymbol{g}|\boldsymbol{f}) = \sum_{m=1}^{M} \left\{-P_{mn} + \frac{g_m}{(\boldsymbol{P}\boldsymbol{f})_m} P_{mn}\right\}. \tag{2.53}$$

Setting $\frac{\partial}{\partial f_n} \log p(\boldsymbol{g}|\boldsymbol{f}) = 0$ yields

$$1 = \frac{1}{\sum_{m=1}^{M} P_{mn}} \sum_{m=1}^{M} \frac{g_m}{(\boldsymbol{P}\boldsymbol{f})_m} P_{mn}. \tag{2.54}$$

Multiplying both side with f and replacing f with a sequence of estimates $\hat{f}^{(k)}$ yields the fixed-point iteration:

$$\hat{f}_n^{(k+1)} = \hat{f}_n^{(k)} \frac{1}{\sum_{m=1}^{M} P_{mn}} \sum_{m=1}^{M} \frac{g_m}{(\boldsymbol{P}\hat{\boldsymbol{f}}^{(k)})_m} P_{mn}. \tag{2.55}$$

The above equation gives MLEM update [45]. In practice, instead of using MLEM, ordered-subset expectation maximization (OSEM) is commonly used for PET reconstruction, which is an accelerated version of MLEM [125]. In OSEM, sinogram data is divided into subsets and these subsets are used to update reconstruction image. The acceleration is approximately proportional to the number of subsets.

In order to control the noise level in reconstructed image, early termination can be applied for MLEM/OSEM algorithm [126]. An alternative way is to incorporate *prior* information such as image smoothness. To do this, we use Bayes' rule to compute the *posterior probability* of image $\boldsymbol{f}$ given observation data $\boldsymbol{g}$. The logarithm of the posterior probability is

$$\begin{aligned}\log p(\boldsymbol{f}\,|\boldsymbol{g}) &= \log \frac{p(\boldsymbol{g}|\boldsymbol{f}\,)p(\boldsymbol{f}\,)}{p(\boldsymbol{g})} \\ &= \log p(\boldsymbol{g}|\boldsymbol{f}\,) + \log p(\boldsymbol{f}\,) - \log p(\boldsymbol{g}).\end{aligned} \tag{2.56}$$

Maximizing equation (2.56) gives *maximum a posterior* (MAP) estimation. In practice, *Gibbs prior* $p(\boldsymbol{f}\,) = \frac{1}{Z}\exp(-\lambda R(\boldsymbol{f}\,))$ is usually chosen as the prior probability, where $R(\boldsymbol{f}\,)$ is a cost function and λ is a tuning parameter [127]. The MAP estimation can be written as

$$\hat{\boldsymbol{f}} = \arg\max_{\boldsymbol{f}} \log p(\boldsymbol{g}|\boldsymbol{f}\,) - \lambda R(\boldsymbol{f}\,). \tag{2.57}$$

To solve this optimization problem, similar to the case for MLEM, we first compute the first order derivative of $\log p(\boldsymbol{f}\,|\boldsymbol{g})$. This gives

$$\frac{\partial}{\partial f_n} \log p(\boldsymbol{f}\,|\boldsymbol{g}) = \sum_{m=1}^{M} \left\{ -P_{mn} + \frac{g_m}{(\boldsymbol{Pf}\,)_m} P_{mn} \right\} - \lambda \frac{\partial R(\boldsymbol{f}\,)}{\partial f_n}. \tag{2.58}$$

By setting $\frac{\partial}{\partial f_n} \log p(\boldsymbol{f}\,|\boldsymbol{g}) = 0$, rearranging the equation, multiplying both side with f, and replacing f with a sequence of $f^{(k)}$, we obtain

$$\hat{f}_n^{(k+1)} = \hat{f}_n^{(k)} \frac{1}{\sum_{m=1}^{M} P_{mn} + \lambda \frac{\partial R(\hat{\boldsymbol{f}}^{(k)})}{\partial f_n}} \sum_{m=1}^{M} \frac{g_m}{(\boldsymbol{P}\hat{\boldsymbol{f}}^{(k)})_m} P_{mn}. \tag{2.59}$$

Notice derivative of $R(\boldsymbol{f}\,)$ in the right side of the above equation is always estimated based on current step $\boldsymbol{f}^{\,(k)}$, rather than the new step $\boldsymbol{f}^{\,(k+1)}$. Thus, this method is known as *one-step late* (OSL) algorithm [128, 129]. Although OSL algorithm is relatively easy to implement and can be applied for any differentiable penalty functions, it is known to be unstable and divergent when strong penalty is applied (i.e. λ is large) [130]. Other methods have also been developed to solve optimization problem in (2.57), such as modified MLEM [131, 132], separable paraboloidal surrogates (SPS) [133], and gradient-based method [134, 135].

2.2.4 Image degradation factors

So far, we only considered an ideal PET system. In practice, PET suffers from different degradation factors. These factors include 1) attenuation, 2) scattered events,

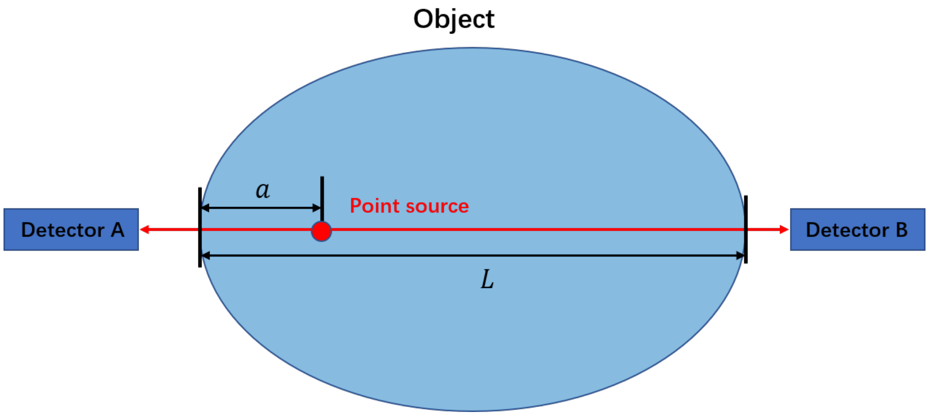


Figure 2-8. Illustration of attenuation in PET.

3) random events, 4) nonuniformity, 5) detector dead time, 6) inter-crystal blurring, 7) positron physics, 8) decay correction, and 9) patient motion.

2.2.4.1 Attenuation

Similar as the case in optical imaging we discussed in 2.1.2.2, photons propagating through the medium may experience photoelectric absorption or be scattered away from its current direction due to scattering. Either of these factors contributes to attenuation along its current propagation direction. For 511 keV γ photons in PET imaging, Compton scattering plays a very dominant role for photon attenuation [94] (unlike the photoelectric effect which increasingly contributes at lower energies). As shown in Fig. 2-8, both photons generated from positron annihilation need to be detected in order to declare an event. Due to attenuation, the probability for each detector to detect a photon can be described using *Beer Lamber's law*. Then the probability of detecting both photons is

$$p(det) = \exp\left(-\int_a \mu(x)dx\right) \times \exp\left(-\int_{L-a} \mu(x)dx\right) = \exp\left(-\int_L \mu(x)dx\right), \quad (2.60)$$

where μ is linear attenuation coefficient. We see that the attenuation along an LOR is independent of the source position along this LOR. This makes attenuation correction more straightforward in PET compared to SPECT. For PET-CT or PET-MR devices,

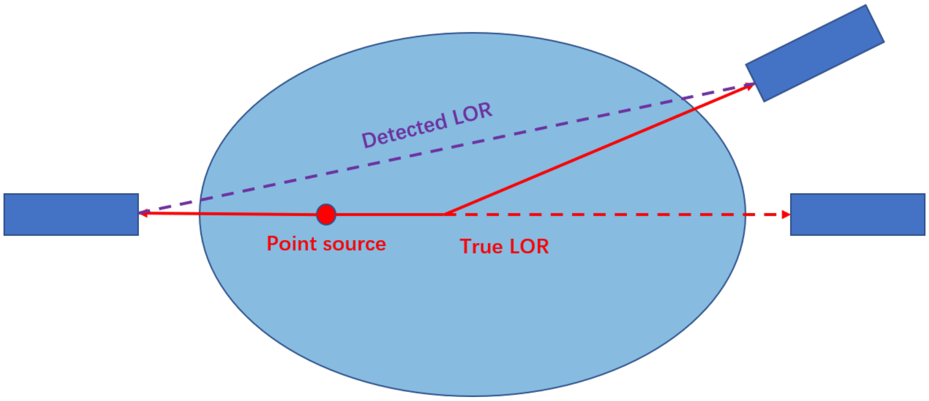


Figure 2-9. Example of scattered event.

attenuation correction can be performed with anatomical information provided by CT or MR systems [120, 121]. For PET-only system, transmission scan can be used for attenuation correction [136]. In this method, images are obtained using an external source with (transmission scan) and without (blank scan) patient. Attenuation correction factors can be computed by dividing the blank scan with the transmission scan.

2.2.4.2 Scattered events

For photon experiencing Compton scattering, although it is considered as attenuated along its original LOR, it may still be detected and declared as an event, which generate a new LOR, as shown in Fig. 2-9. Such event is known as scattered event. Scattered events cause mispositioning of LOR, which reduces image contrast. Scattered events can constitute 10-20% of all events for 2D PET and 40-60% for 3D PET [95]. Scatter correction can be achieved with fitting technique [137], energy window methods [138, 139], or model-based methods [140, 141].

2.2.4.3 Random events

In PET imaging, it is possible that only one photon in an annihilation event is not detected due to its propagation direction, attenuation or simply passing through

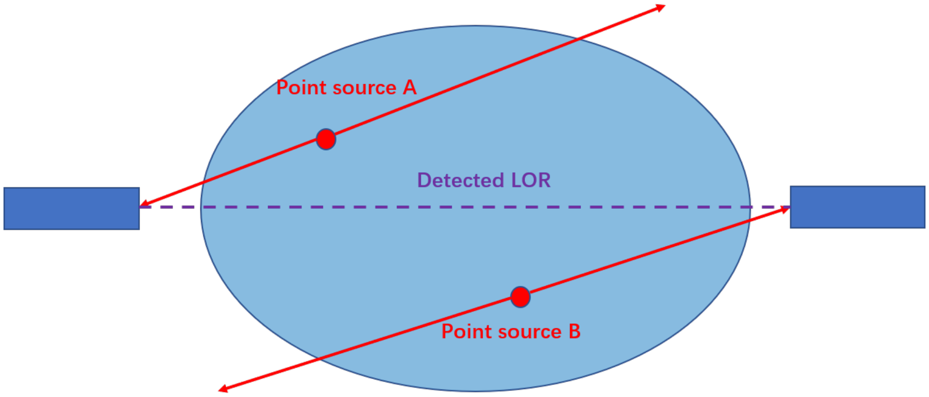


Figure 2-10. Example of random event.

detector without being absorbed, while the other photon is successfully detected. If photon pairs from two such events are detected within the coincidence time window and considered as a single event, this event is known as random event. Fig. 2-10 shows an example of random event. Random events add relatively uniform background across the reconstructed image and thus reduce contrast of reconstructed image. The rate of random event in a detector pair is given by [93]

$$R_{random,AB} = \Delta T R_{single,A} R_{single,B}, \tag{2.61}$$

where ΔT is the coincidence time window, $R_{single,A}$ and R_{single_B} are the rate of single events at detector A and B, respectively. As can be seen from the equation, random rate is proportional to the width of the coincidence time window. On the other hand, since the rate of single events for one detector increases linearly with the total amount of activity, random rate increases as the square of the total activity. Since random rate is only related to the width of the time window, but independent of time point where the window is set, a delayed time window can be used to obtain the number of random event and used for correction. An alternative way is to measure counting rate of individual detectors and compute random rate with equation (2.61).

2.2.4.4 Nonuniformity

Efficiency of LORs joining detectors varies due to influence from detectors as well as geometry of PET scanner [142]. For individual detectors, sensitivity changes due to small differences in, for example, crystal dimensions, efficiency and PMTs gains. In addition, geometry influences include i) different photon incident angles for LORs at different detectors, which changes effective thickness as well as cross section of crystals, and ii) varying efficiency of LORs across different detector rings [143]. These factors result in nonuniformities in PET scanning, with undesirable artifacts in reconstructed images. Correction for this is known as normalization. The normalization factors can be obtained by performing a normalization scan for a uniform or rotating radioactive source.

2.2.4.5 Detector dead time

Dead time related to the time required by a detector to process individual detected events. During this time the detector is unable to process incoming events, which results in loss of count. The loss due to dead time increases as total amount of activity increases. This limits the amount of injected radioactivity. Different components of a PET system could contribute to the dead time. A contribution comes from detector integration time, which is related to crystal decay time [144]. If a photon deposits energy in the detector crystal while charge from previous event is still being integrated, this gives the pulse pile-up effect. The pulse pile-up effect leads to two results: either collected charge exceeds the upper threshold of energy window and is rejected, or the two events are treated as one. Both cases result in count loss. In addition, the detector electronics usually have a reset time, during which the sub-system is unable to accept further events. In coincidence detection circuit, it is also possible that more than two events might occur during the coincidence time window, which is known as multiple coincidence. In this case, events comprising the multiple coincidence will be rejected.

Also, processing a coincidence event takes time, during which new coincidences event cannot be accepted. In PET scanners, data channels are multiplexed into smaller shared circuits, which has higher data rate and contribute to the overall system dead time [145].

The counting systems can be classified as paralyzable and nonparalyzable systems. In a nonparalyzable system, if an event occurs during the dead time of a preceding event, the new event will simply be ignored and have no influence on subsequently occurring events. On the other hand, for a paralyzable system, each event introduces a dead time whether or not that event is actually counted. This makes a new event extend the dead time of a previous event and thus further influence subsequent events.

Deadtime correction can be performed by scaling the observed counts by correction factors that are measured from singles rate [94].

2.2.4.6 Inter-crystal blurring

An incoming γ photon may enter one detector crystal but is detected in another crystal. When this happens, it produces mispositioned LOR and gives so-called inter-crystal blurring that degrades image resolution [93]. Two cases may cause inter-crystal blurring. (i) when entering a detector crystal, γ photon may experience Compton scattering and is deflected into another crystal, as shown in Fig. 2-11 (a). This is known as *inter-crystal scattering.* (ii) If a photon enters crystal at some angle, it may penetrate through current detector and is detected in the neighboring crystal, which is known as *inter-crystal penetration*, as shown in Fig. 2-11 (b).

This latter inter-crystal penetration effect leads to the so-called parallax effect, which refers to the phenomenon that the apparent width of a detector element increases with increasing radial offset in a PET scanner since the interaction depth of the γ photons are unknown. This cause a degradation in radial spatial resolution with increasing distances from the center of the scanner field of view (FOV) [146]. The

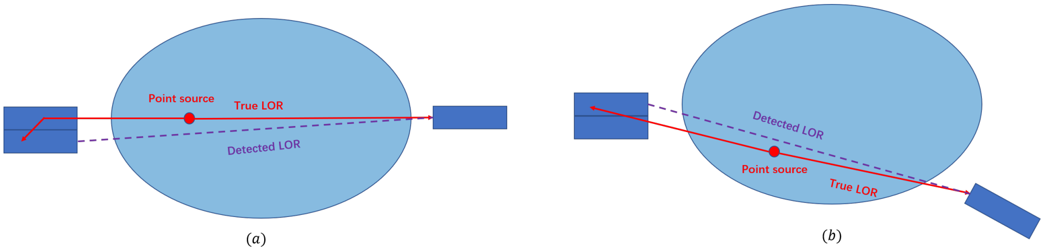


Figure 2-11. Example of (a) inter-crystal scattering and (b) inter-crystal penetration.

parallax effect can be reduced by using depth-of-interaction (DOI) encoding technology [147–150] as well as resolution modeling techniques [151].

2.2.4.7 Positron physics

Two factors from positron physics could degrade spatial resolution of PET scanner. The first one is positron range. As we have discussed in 2.2.1, emitted positron travels a short distance before annihilation, as shown in Fig. 2-4. As a result, detected LOR only covers the location where positron annihilation occurs, but not precisely the location where the decaying radionuclide emitted the positron. We have listed positron ranges for commonly used radionuclide in Table 2-I. We can observe that for some radionuclides, such as ^{82}Rb and ^{68}Ga, their positron ranges can be large and could severely degrade system resolution if they are not corrected. The second factor is photon non-collinearity. In practice, the photon pair generated from annihilation does not travel in exactly antiparallel direction. The angular distribution is approximately Gaussian with full width at half maximum (FWHM) approximately 0.5 degree [94]. This effect is due to small residual momentum of the positron when it annihilates with a electron. Its effect on spatial resolution is linearly dependent on the separation distance of detectors and is described as

$$FWHM = 0.0022 \times D, \tag{2.62}$$

where D is the diameter of the FOV. For whole-body PET scanner, diameter of FOV is about 80 cm, which gives about 2 mm FWHM blurring due to photon non-collinearity

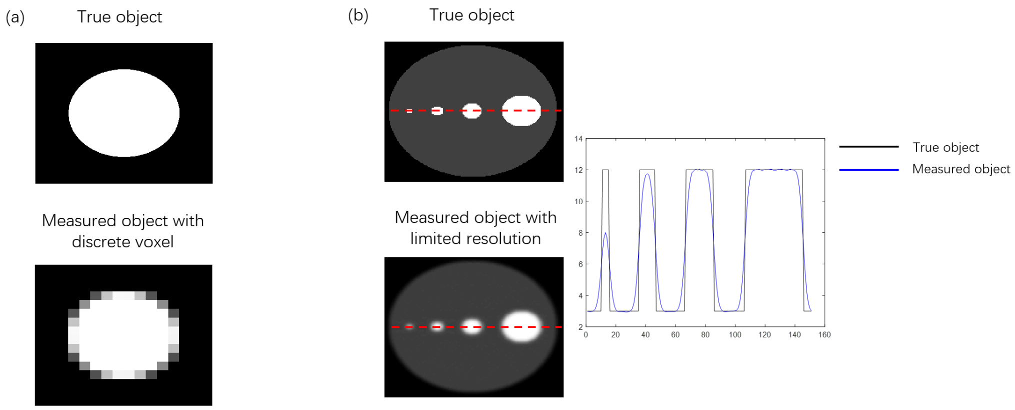


Figure 2-12. Illustration of PVE due to (a) tissue fraction effect and (b) limited spatial resolution

effect.

2.2.4.8 Decay correction

Radiotracers experience radioactive decay as time passes and thus the emitted positrons also decrease. Such decay needs to be corrected in order to get quantitative PET images independent of physical decay. The correction can be performed by scaling the image counts with the decay correction factor computed using decay constant of radionuclides, start and end time points of the PET frame.

2.2.4.9 Patient motion

During PET scan, patient movement could severely degrade image quality such as introducing motion blurring and attenuation mismatch artifacts [152]. Patient motion can be roughly categorized into (i) unwanted body movement, (ii) motion due to cardiac cycles and (iii) motion due to respiratory cycles [153].

2.2.5 Partial volume effect

The spatial resolution of a PET scanner is primarily determined by the size of detectors [154]. Besides, we discussed in the last section several factors, including positron physics and detector inter-crystal blurring, that further degrade the spatial resolution of a PET system. Commercially available PET scanners usually have spatial resolutions 4-6 mm [93]. As a result, the limited spatial resolution leads to the so-called partial volume effect (PVE) [122]. In general, PVE can be defined as the loss in apparent activity that occurs when an object is imaged by an imaging instrument [155, 156]. PVE is attributed to two effects: tissue-fraction effect and spill-over effect [97, 156].

In the tissue-fraction effect, due to the finite voxel size, each individual voxel can contain two or more tissue types. This usually occurs at the boundary between regions of different tissue-type, as shown in Fig. 2-12 (a). When PVE is mentioned in CT or MRI images, it usually refers to this effect [157, 158].

For the spill-over effect, due to the blurring introduced by finite spatial resolution of the imaging system, activity could spill into adjacent regions. The spill-over effect not only causes blurry appearance of PET images, but also leads to quantitative overestimation in cold region and underestimation in hot region. An example of spill-over effect is shown in Fig. 2-12 (b).

Apart from the above two phenomena, patient motion, especially respiratory motion, could also introduce blurring effect that gives additional PVE [159]. Motion correction itself is an active research field. More discussion on motion correction can be found in [152, 153, 160–166].

Structures smaller than 3 times the FWHM of scanner resolution will be influenced by PVE [97, 167]. The main consequence of PVE is the introduction of bias that is related to the size of region of interests (ROIs). This may lead to inaccurate clinical assessment of PET images. For example, for brain imaging, the presence of tissue atrophy could increase PVE. As a result, it is hard to distinguish such apparent loss of radioactivity due to increased PVE from true loss of tissue uptake [168]. In oncologic imaging, PVE influences the apparent tumor size and heterogeneity, which could further influence radiotherapy treatment planning [97, 169]. Thus, in order to obtain accurate quantification and clinical interpretation of PET images, it is important to correct such partial volume effects. In Chapter 5, we will discuss methods to perform partial volume correction (PVC).

Chapter 3

Incorporating reflection boundary conditions in the Neumann-series-based radiative transport equation

In this chapter, we describe a novel Neumann-series RTE that models the reflection of photons at the boundary between tissue and external medium. We use this Neumann series RTE to investigate the improvement in estimating optical coefficients when reflection of photons at the boundary is modeled. For this purpose, we also propose a novel reconstruction technique that incorporates the Neumann-series formalism to estimate the optical coefficients of the tissue. These methods are evaluated using a Monte Carlo simulated three-dimensional (3D) DOI imaging system. It was observed that the average root-mean-square error (RMSE) for the output images and the estimated absorption coefficients reduced by 38% and 84%, respectively, when the reflection boundary conditions were incorporated. These results demonstrate the importance of incorporating boundary conditions that model the reflection of photons at the tissue-external medium interface.

3.1 Introduction

Photon propagation modeling plays an important role in biomedical optics applications such as diffuse optical imaging (DOI), diffuse optical spectroscopy, FMT, and light dosimetry modeling [51, 69, 170–173]. The RTE is a well-known method for modeling this light propagation. Due to the computational complexity of solving the RTE directly, different methods have been proposed to provide approximation solution to the RTE [47, 52, 56, 62]. Among these methods, the DA is widely used due to its simplicity and fast implementation in complicated geometry [48]. However, as we discussed in 2.1.2.3, the DA assumes isotropic source as well as strong scattering medium, which is not correct near tissue surface, in anisotropic tissues, and in regions of high absorption or low scatter [50, 53] As a result, it cannot accurately describe light propagation in highly absorbing regions such as haematomas, void-like spaces such as ventricles and the subarachnoid-space, and for small tissue geometries, such as whole-body imaging of small animals [53, 174, 175].

We discussed a Neumann-series-based RTE method in 2.1.2.3, which solves the RTE in using integral form equation. This method requires solving only a single equation. Further, the method also provided improved accuracy for imaging setups where the DA-based methods have limitations [54]. Additionally, the method provided a novel intuitive framework to physically interpret the RTE solution. Each term in the Neumann-series RTE describes contributions from photons that have scattered a specific number of times. Thus, the method can numerically quantify the contribution from photons that, for example, do not scatter (ballistic photons), or scatter only a certain number of times. However, the existing Neumann-series method inaccurately assumes that there is no reflection of photons at the boundary. On the other hand, practical biomedical optics applications, as we mentioned above, usually require placing detectors on the surface of the tissue, which leads to a refractive index mismatch when photons exiting from the tissue to the external medium. Accounting for the boundary conditions arising due to this refractive index mismatch is necessary for accurate modeling of photon propagation. It has been observed that not accounting for boundary conditions accurately leads to erroneous modeling of scattered light [176], and 50% or more errors in estimating the optical coefficients of the tissue [177].

Studies on incorporating boundary conditions while modeling photon transport in biological tissue have been conducted. For example, Marshak-type boundary condition has been proposed for analytical solution of the RTE. This condition incorporates reflection of photons accurately using Fresnel's equations [178–180]. For DA, different boundary conditions including partial-current, extrapolated, and vacuum boundary conditions have been proposed. These conditions have varying degrees of inaccuracy [177, 181, 182]. A corrected DA method has been proposed to correct these inaccuracies [183, 184].

In the following, we describe how to incorporate reflection boundary condition into Neumann-series-based RTE model. In addition, we also present a reconstruction technique for DOI using Neumann-series-based RTE as forward model.

3.2 Methods

3.2.1 The Neumann-series RTE modeling the boundary conditions

We have introduced the RTE in 2.1.2.2. Here we rewrite the steady-state RTE equation as

$$\hat{\boldsymbol{s}} \cdot \nabla w(\boldsymbol{r}, \hat{\boldsymbol{s}}) + \mu_{tot} w(\boldsymbol{r}, \hat{\boldsymbol{s}}) = \frac{1}{c_m} \left[\Xi(\boldsymbol{r}, \hat{\boldsymbol{s}}) + \mu_s(\boldsymbol{r}) \int_{4\pi} d\Omega' p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}'; \boldsymbol{r}) w(\boldsymbol{r}, \hat{\boldsymbol{s}}') \right], \qquad (3.1)$$

where $w(\boldsymbol{r}, \hat{\boldsymbol{s}})$ is the photon distribution function at location $\boldsymbol{r} = (x, y, z)$ in direction $\hat{\boldsymbol{s}}$, $\mu_{tot}(\boldsymbol{r}) = \mu_a(\boldsymbol{r}) + \mu_s(\boldsymbol{r})$ is the total attenuation coefficient, $\mu_a(\boldsymbol{r})$ and $\mu_s(\boldsymbol{r})$ are absorption and scattering coefficients, respectively, c_m is the speed of light in the medium,

$\Xi(\boldsymbol{r}, \hat{\boldsymbol{s}})$ denotes a monochromatic mono-energetic emission source, and $p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}'; \boldsymbol{r})$ is Henyey-Greenstein scattering phase function and is defined as

$$p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}'; \boldsymbol{r}) = \frac{1}{4\pi} \frac{1 - g^2}{[1 + g^2 - 2g(\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{s}}')]^{3/2}}, \tag{3.2}$$

where $g \in [-1, 1]$ is anisotropy factor. The RTE can be represented in an integral form as

$$w = \boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}w, \tag{3.3}$$

where $\boldsymbol{\mathcal{K}}$ and $\boldsymbol{\mathcal{X}}$ are scattering and attenuation operators, respectively. The effect of the scattering operator on the distribution function is

$$[\boldsymbol{\mathcal{K}}w](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \int_{4\pi} d\Omega' c_m \mu_s(\boldsymbol{r}) p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}'; \boldsymbol{r}) w(\boldsymbol{r}, \hat{\boldsymbol{s}}'). \tag{3.4}$$

Similarly, the attenuation operator performs as

$$[\boldsymbol{\mathcal{X}}w](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \frac{1}{c_m(\boldsymbol{r})} \int_0^\infty d\lambda w(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda, \hat{\boldsymbol{s}}) \exp\left[-\int_0^\lambda d\lambda' \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda')\right], \tag{3.5}$$

where λ corresponds to the length variable. A schematic demonstrating the scattering and attenuation operators for the 3D DOI system considered in this chapter is given in Fig. 3-1. Equation (3.3) can be alternatively written in a Neumann-series form as follows:

$$w = \boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi + \ldots \tag{3.6}$$

The equation has a very intuitive physical interpretation; the contribution of photons that have scattered n times is given by the term $\boldsymbol{\mathcal{X}}(\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}})^n\Xi$ in this series.

To incorporate the boundary conditions into the Neumann series, we start with a first principles treatment of light propagation in tissue. Similar to Schweigher et al. [185], we define the boundary of the tissue to be an infinitesimally thin layer where only the reflection event occurs. This thin boundary layer is illustrated in Fig. 3-2. Due to the refractive index mismatch, photons incident on the boundary are reflected within this thin layer. Therefore, the thin layer acts as a source of photon emission.

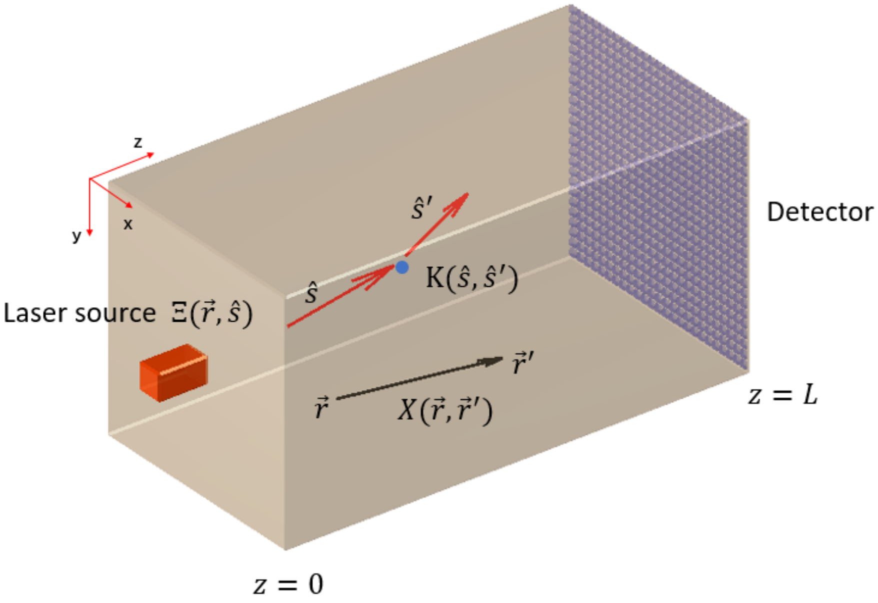


Figure 3-1. Physical interpretation of the various operators in the RTE in considered DOI setup.

Let $\boldsymbol{\mathcal{R}}$ denote the reflection operator. An alternative way to think about the reflection operation is to consider it as a scattering operation, except that the scattering phase function is given by the laws of reflection. Either of these two interpretations, when modeled in equation (3.3), leads to the following form for the RTE:

$$w = \boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}w + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}w. \tag{3.7}$$

The operator $\boldsymbol{\mathcal{R}}$ is defined using Fresnel's laws of reflection. Let $\hat{\boldsymbol{s}}$, $\hat{\boldsymbol{s}}'$ and $\hat{\boldsymbol{n}}$ denote the angle of incidence, reflection, and normal to the boundary surface. In accordance with Fresnel's laws of reflection,

$$[\boldsymbol{\mathcal{R}}w](\boldsymbol{r},\hat{\boldsymbol{s}}) = c_m \int_{\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{s}}'>0} d\Omega' R(\hat{\boldsymbol{s}}')\delta[\hat{\boldsymbol{s}} - \hat{\boldsymbol{s}}' + 2(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{s}}')\hat{\boldsymbol{n}}]w(\boldsymbol{r},\hat{\boldsymbol{s}}')\delta(b(\boldsymbol{r})), \tag{3.8}$$

where $\delta(\boldsymbol{r})$ is the delta function and $b(\boldsymbol{r}) = 0$ is the equation of the boundary. Thus, $\delta(b(\boldsymbol{r}))$ denotes that the reflection operator acts only on photons in the infinitesimally thin boundary layer. For example, if the boundary is a plane normal to $\hat{\boldsymbol{n}}$ and at a perpendicular distance p from the origin, then $b(\boldsymbol{r}) = \boldsymbol{r}\cdot\hat{\boldsymbol{n}} - p$. Further, $R(\hat{\boldsymbol{s}}')$ is

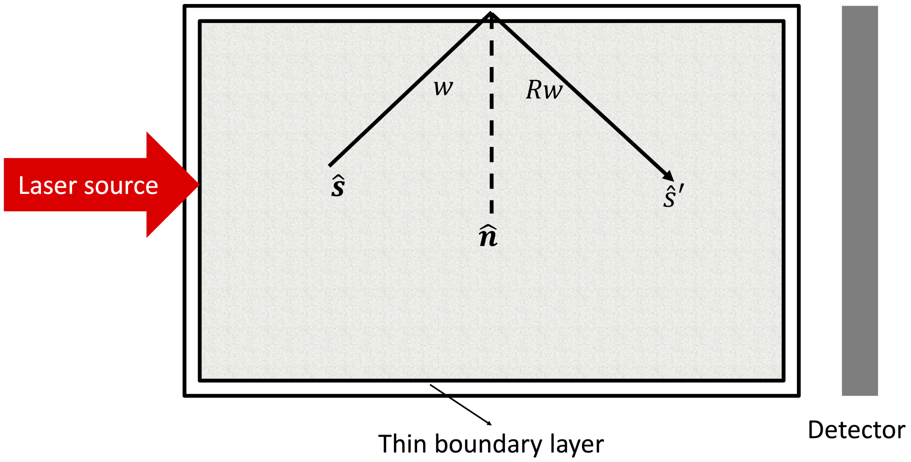


Figure 3-2. A schematic describing the thin boundary layer in the considered DOI setup.

reflectivity coefficient. Let n_i and n_t denote refractive index of incident medium and transmitted medium, respectively. Also let θ_i, θ_t and θ_c denote incident, transmitted, and critical angle, respectively. Then we have

$$R(\hat{\boldsymbol{s}}) = \begin{cases} \frac{1}{2}\left(\frac{n_i\cos\theta_i - n_t\cos\theta_t}{n_i\cos\theta_i + n_t\cos\theta_t}\right)^2 + \frac{1}{2}\left(\frac{n_t\cos\theta_i - n_i\cos\theta_t}{n_t\cos\theta_i + n_i\cos\theta_t}\right)^2, \text{for } \theta_i < \theta_c. \\ 1, \text{for } \theta_i > \theta_c. \end{cases} \tag{3.9}$$

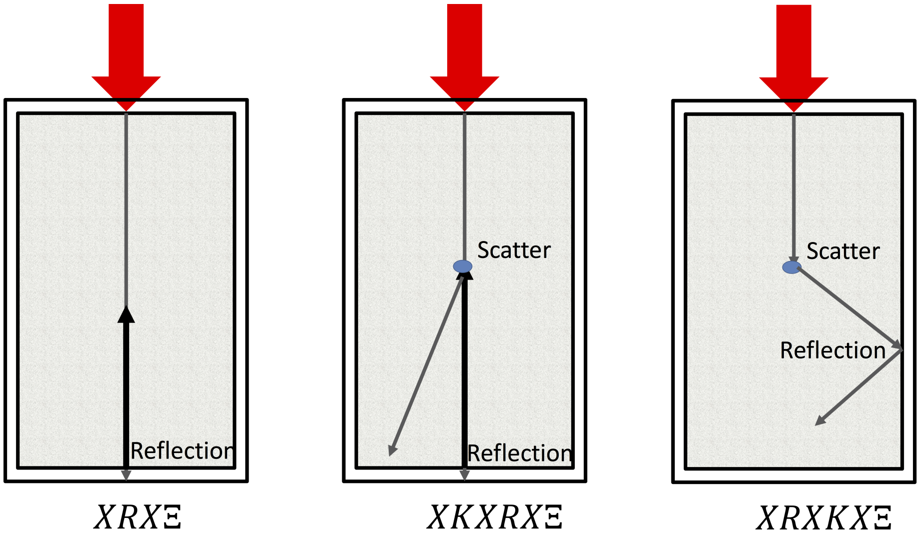
Taking the terms involving the distribution function on the left side of equation (3.7) yields

$$[\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}} - \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}]w = \boldsymbol{\mathcal{X}}\Xi. \tag{3.10}$$

A Neumann-series solution for the above equation is

$$\begin{aligned} w &= [\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}} - \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}]^{-1}\boldsymbol{\mathcal{X}}\Xi \\ &= [\boldsymbol{\mathcal{I}} + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}} + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}} + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}} + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}} + \cdots]\boldsymbol{\mathcal{X}}\Xi, \end{aligned} \tag{3.11}$$

Similar to the originally proposed Neumann-series RTE, the various terms in the above expansion have physical interpretations, as shown in Fig. 3-3. For example, the term $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi$ represent the photons that are reflected at the boundary and subsequently transmitted back into the medium. The photons that reflects back into the medium and then get scattered are represented by the term $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi$. The term $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi$ denotes photons that are scattered in the medium and subsequently reflected from a boundary surface.



$XRX\Xi$ $XKXRX\Xi$ $XRXKX\Xi$

Figure 3-3. Physical interpretation of the three reflection terms in the RTE.

3.2.2 Implementation

The Neumann-series formalism was implemented for a 3-D DOI setup as shown in Fig. 3-1. The output image was obtained by a pixellated contact detector. For computational reasons, the effect of reflection was accounted using only three terms that contained the reflection operator, as follows:

$$\begin{aligned} w = {} & \boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi + \\ & \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi + \cdots \qquad (3.12) \\ & = \text{Original Neumann series} + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi. \end{aligned}$$

The original Neumann series has been implemented in [54]. For the three additional terms involving reflection operator, we will show their mathematical expressions below. The detail derivations of these terms can be found in Appendix 3.A.

The first term $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi$ describes photons that are emitted from the source, pass through the medium, incident on the exit boundary, and then reflected back into the medium. In our DOI setup, the laser source emits a unidirectional beam along the optical axis, considered as the z axis. Denote the plane where the laser source is incident by $z = 0$ and the length of the medium along the z axis by L. Denote the profile of the beam along the (x, y) coordinates by $h(x, y)$ and let α denote the intensity of the laser source. Then the source term is

$$\Xi(\boldsymbol{r}, \hat{\boldsymbol{s}}) = \alpha h(x, y)\delta(\hat{\boldsymbol{s}} - \hat{\boldsymbol{z}})\delta(z). \tag{3.13}$$

For this source term, the expression for the $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi$ term, as derived in Appendix 3.B, is

$$[\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \frac{\alpha R(\hat{\boldsymbol{z}})}{c_m} h(x, y)\delta(\hat{\boldsymbol{s}}+\hat{\boldsymbol{z}}) \exp\left[-\int_0^L d\lambda' \mu_{tot}(x, y, L - \lambda') - \int_0^{L-z} d\lambda' \mu_{tot}(\boldsymbol{r} + \hat{\boldsymbol{z}}\lambda')\right]. \tag{3.14}$$

The second term $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi$ describes photons that scatter once after reflection before reaching the detector surface. To compute this term, we treat the term $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi$, given by equation (3.14), as a new source term Ξ'. We apply scattering operator to this source term and obtain

$$[\boldsymbol{\mathcal{K}}\Xi'](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \alpha\mu_s R(\hat{\boldsymbol{z}})h(x, y)p(\hat{\boldsymbol{s}}, -\hat{\boldsymbol{z}}) \exp\left[-\int_0^L d\lambda' \mu_{tot}(x, y, L - \hat{\boldsymbol{z}}\lambda') - \int_0^{L-z} d\lambda' \mu_{tot}(\boldsymbol{r} + \hat{\boldsymbol{z}}\lambda')\right]. \tag{3.15}$$

Implementing this term in Cartesian coordinate can be difficult. As we discussed in 2.1.2.2, spherical harmonics can be used to expand this term based on its angular distribution. If we express the term $[\boldsymbol{\mathcal{K}}\Xi'](\boldsymbol{r}, \hat{\boldsymbol{s}})$ in spherical harmonics, we obtain

$$[\boldsymbol{\mathcal{D}}\chi']_{lm}(\boldsymbol{r}) = \alpha\mu_s R(\hat{\boldsymbol{z}})h(x, y)\sqrt{\frac{2l+1}{4\pi}} g^l \exp\left[-\int_0^L d\lambda' \mu_{tot}(x, y, L - \hat{\boldsymbol{z}}\lambda') - \int_0^{L-z} d\lambda' \mu_{tot}(\boldsymbol{r} + \hat{\boldsymbol{z}}\lambda')\right]. \tag{3.16}$$

The attenuation operator is then applied numerically on this term. The expression for the attenuation operator in the spherical harmonic basis has been introduced in 2.1.2.3.

The third term $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi$ describes photons that scatter once before being reflected at boundary. This boundary is the set of all planes surrounding the medium, as shown in Fig. 3-2, and is thus defined as $\sum_i \delta(\boldsymbol{r} \cdot \hat{\boldsymbol{n}}_i - p_i)$, where the index i denotes the different planes, $\hat{\boldsymbol{n}}_i$ denotes the unit normal vector to the i^{th} plane and p_i denotes the distance of the i^{th} plane from the origin along the normal vector. With this notation, the expression for the $\boldsymbol{\mathcal{XRXKX}}\Xi$ term, as derived in Appendix 3.B, is

$$[\boldsymbol{\mathcal{XRXKX}}\Xi](\boldsymbol{r}, \hat{\boldsymbol{s}}') = \alpha \sum_i \int_\infty d^3r' \frac{1}{|\boldsymbol{r} - \hat{\boldsymbol{s}}'\lambda_{21i} - \boldsymbol{r}'|^2} \frac{1}{\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}}_i} R(\hat{\boldsymbol{s}}_{10}) \delta(\hat{\boldsymbol{s}}' - \hat{\boldsymbol{s}}_{10} + 2(\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{s}}_{10})\hat{\boldsymbol{n}}) \times \\ h(x', y') p(\hat{\boldsymbol{s}}_{10}, \hat{\boldsymbol{z}}) \exp\left[- \int_0^{z'} d\lambda' \mu_{tot}(\boldsymbol{r}' - \hat{\boldsymbol{z}}\lambda') - \int_0^{\lambda} d\lambda' \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda_{21i} - \hat{\boldsymbol{s}}_{10}\lambda') \right], \tag{3.17}$$

where

$$\lambda_{21i} = \frac{\boldsymbol{r} \cdot \hat{\boldsymbol{n}}_i - p_i}{\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}}_i}, \tag{3.18}$$

and

$$\boldsymbol{s}_{10} = \frac{\boldsymbol{r} - \boldsymbol{r}'}{|\boldsymbol{r} - \boldsymbol{r}'|}. \tag{3.19}$$

From the computed distribution function, the transmitted flux on the detector face for the DOI system is computed as follows. Denote the x and y dimensions of each pixel by Δx and Δy, respectively. Denote the transmitted flux detected by the m^{th} pixel with center at $\boldsymbol{r}_m$ by Φ_m. Assuming that the distribution function over a pixel is approximately constant, we get

$$\Phi_m = c_m \Delta x \Delta y \int_{2\pi} d\Omega T(\hat{\boldsymbol{s}})(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{s}}) w(\boldsymbol{r}_m, \hat{\boldsymbol{s}}), \tag{3.20}$$

where $T(\hat{\boldsymbol{s}}) = 1 - R(\hat{\boldsymbol{s}})$ is transmission coefficient. As an example, the flux due to the $\boldsymbol{\mathcal{XRXKX}}\Xi$ term is derived in Appendix 3.B.

3.2.3 Neumann-series-based reconstruction algorithm for DOI with boundary conditions

Denote the absorption/scattering coefficient at location $\boldsymbol{r}$ by the function $\mu(\boldsymbol{r})$. The function is discretized using the spatial basis function $\phi_n(\boldsymbol{r})$, as below:

$$\mu(\boldsymbol{r}) = \sum_{n=1}^{N} \mu_n \phi_n(\boldsymbol{r}), \tag{3.21}$$

where μ_n denote the optical coefficient (scattering or absorption) for the n^{th} basis function. The spatial basis functions could be the commonly used voxels, but the method is generalizable to other spatial basis functions too. Denote the N-dimensional vector of the coefficients μ_n by μ. Denote the image acquired by the pixelated detector in the DOI setup by the M-dimensional vector $\boldsymbol{g}$. Our objective is to estimate $\boldsymbol{\mu}$ given $\boldsymbol{g}$. For this purpose, we derive a gradient-descent-based approach.

We denote the mean noiseless image as a function of the scattering and absorption coefficient by $\bar{\boldsymbol{g}}(\boldsymbol{\mu})$. In our reconstruction approach, the objective is to estimate the value of $\boldsymbol{\mu}$ that minimizes the L_2 norm of the error between the measured data $\boldsymbol{g}$ and the $\bar{\boldsymbol{g}}(\boldsymbol{\mu})$. Mathematically, the objective is to estimate $\hat{\boldsymbol{\mu}}$ that satisfies

$$\hat{\boldsymbol{\mu}} = \arg\min_{\boldsymbol{\mu}} \boldsymbol{\Psi}(\boldsymbol{\mu}), \tag{3.22}$$

where $\boldsymbol{\Psi}(\boldsymbol{\mu}) = \|\boldsymbol{g} - \bar{\boldsymbol{g}}(\mu)\|_2^2$, is the objective function, and $\|\boldsymbol{x}\|_2^2$ denotes the square of the L_2 norm of the vector $\boldsymbol{x}$. To implement the gradient-descent method, we need to calculate the derivative of $\boldsymbol{\Psi}(\boldsymbol{\mu})$, which can be written as

$$\nabla\boldsymbol{\Psi}(\boldsymbol{\mu}) = -2\left(\frac{d\bar{\boldsymbol{g}}(\boldsymbol{\mu})}{d\boldsymbol{\mu}}\right)^T (\boldsymbol{g} - \bar{\boldsymbol{g}}(\boldsymbol{\mu})). \tag{3.23}$$

The calculation of the above equation requires computing the derivative of $\bar{\boldsymbol{g}}(\boldsymbol{\mu})$ with respect to μ_a and μ_s. Let $h_m(\boldsymbol{r}, \hat{\boldsymbol{s}})$ denote the detector response function of the m^{th} detector pixel. Then, the m^{th} component of $\bar{\boldsymbol{g}}(\boldsymbol{\mu})$, denoted by $\bar{g}_m$, is

$$\bar{g}_m = \langle h_m, w\rangle, \tag{3.24}$$

where $\langle\cdot\ ,\cdot\rangle$ represents inner product. Taking the derivative on both sides with respect to μ_n yields

$$\frac{\partial\bar{g}_m}{\partial\mu_n} = \left\langle h_m, \frac{\partial w}{\partial\mu_n}\right\rangle. \tag{3.25}$$

From equation (3.7), we have

$$\frac{\partial w}{\partial\mu_n} = \frac{\partial\boldsymbol{\mathcal{X}}}{\partial\mu_n}\Xi + \frac{\partial\boldsymbol{\mathcal{X}}}{\partial\mu_n}\boldsymbol{\mathcal{R}}w + \frac{\partial\boldsymbol{\mathcal{X}}}{\partial\mu_n}\boldsymbol{\mathcal{K}}w + \boldsymbol{\mathcal{X}}\frac{\partial\boldsymbol{\mathcal{K}}}{\partial\mu_n}w + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\frac{\partial w}{\partial\mu_n} + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\frac{\partial w}{\partial\mu_n}. \tag{3.26}$$

As derived in [186], terms containing partial derivatives of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{K}}$ are

$$\frac{\partial\boldsymbol{\mathcal{X}}}{\partial\mu_n}\Xi = -c_m\boldsymbol{\mathcal{X}}(\phi_n\boldsymbol{\mathcal{X}}\Xi), \tag{3.27}$$

$$\frac{\partial \boldsymbol{\mathcal{X}}}{\partial \mu_n} \boldsymbol{\mathcal{K}} w = -c_m \boldsymbol{\mathcal{X}}(\phi_n \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{K}} w), \tag{3.28}$$

and

$$\frac{\partial \boldsymbol{\mathcal{K}}}{\partial \mu_n} w = \epsilon \phi_n \boldsymbol{\mathcal{K}}_1 w, \tag{3.29}$$

where $\epsilon = 1$ for $\mu = \mu_s$ and $\epsilon = 0$ for $\mu = \mu_a$, and the operator $\boldsymbol{\mathcal{K}}_1$ is defined as

$$[\boldsymbol{\mathcal{K}}_1 w](\boldsymbol{r}, \hat{\boldsymbol{s}}) = c_m \int d\Omega' p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{s}}') w(\boldsymbol{r}, \hat{\boldsymbol{s}}'). \tag{3.30}$$

Further, as derived in Appendix 3.B

$$\frac{\partial \boldsymbol{\mathcal{X}}}{\partial \mu_n} \boldsymbol{\mathcal{R}} w(\boldsymbol{r}, \hat{\boldsymbol{s}}) = -c_m \boldsymbol{\mathcal{X}}(\phi_n \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{R}} w). \tag{3.31}$$

Substituting the expressions from (3.27)-(3.31) in equation (3.26), we have

$$\frac{\partial w}{\partial \mu_n} = c_m \boldsymbol{\mathcal{X}} \phi_n (\boldsymbol{\mathcal{X}} \Xi + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{R}} w + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{K}} w) + \boldsymbol{\mathcal{X}} \epsilon \phi_n \boldsymbol{\mathcal{K}}_1 w + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{R}} \frac{\partial w}{\partial \mu_n} + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{K}} \frac{\partial w}{\partial \mu_n}. \tag{3.32}$$

As indicated in equation (3.7), $\boldsymbol{\mathcal{X}} \Xi + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{R}} w + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{K}} w = w$. Thus

$$\frac{\partial w}{\partial \mu_n} = -c_m \boldsymbol{\mathcal{X}} \phi_n w + \boldsymbol{\mathcal{X}} \epsilon \phi_n \boldsymbol{\mathcal{K}}_1 w + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{R}} \frac{\partial w}{\partial \mu_n} + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{K}} \frac{\partial w}{\partial \mu_n}. \tag{3.33}$$

Define $S = -c_m \phi_n w + \epsilon \phi_n \boldsymbol{\mathcal{K}}_1 w$. The above equation can then be rewritten as

$$\frac{\partial w}{\partial \mu_n} = \boldsymbol{\mathcal{X}} S + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{R}} \frac{\partial w}{\partial \mu_n} + \boldsymbol{\mathcal{X}} \boldsymbol{\mathcal{K}} \frac{\partial w}{\partial \mu_n}, \tag{3.34}$$

Comparison to equation (3.7) reveals that the expression for the gradient is the same as the original RTE, but with a different source term. Therefore, the gradient of the distribution function is computed by simply executing the Neumann-series RTE but with the source term S.

From the gradient of the distribution function and using equation (3.25), the gradient of g_m with respect to μ_n is obtained. Using a simple iterative gradient-descent approach, the computed gradient is used to update the values of the optical coefficients in each iteration until convergence is achieved. Thus, the absorption and scattering coefficients are estimated. Specifically, for a homogeneous medium, the N-dimensional coefficients vector $\boldsymbol{\mu}$ in the above derivation becomes a scalar value μ. The calculation of $\frac{\partial \bar{g}_m}{\partial \mu}$ remains the same as the above derivation.

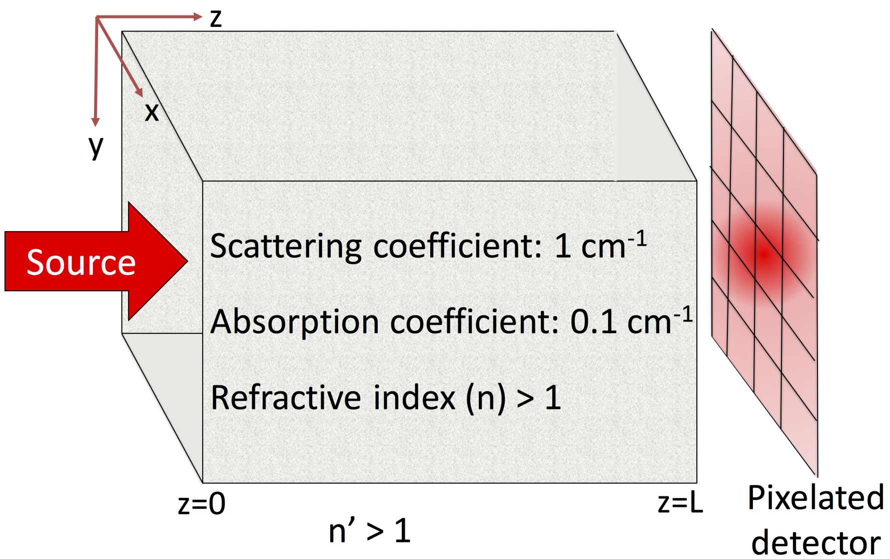


Figure 3-4. The experimental setup simulated to evaluate the proposed framework.

3.2.4 Simulation experimental setup

The proposed Neumann-series RTE approach was validated for a 3D DOI setup (Fig. 3-4). The scattering medium was a cube with each side of length 2 cm. The scattering and absorption coefficients in the medium were 1 cm^{-1} and 0.01 cm^{-1}. The medium was small geometry, had a relatively low scattering coefficient, and a collimated source. The choice for this setup was made to study the performance of the proposed method in a scenario where the diffusion approximation-based methods are known to have limitations. In the simulation setup, across different experiments, the refractive index of the media was varied from 1.1 to 1.5, while the refractive index of the external medium was kept as 1. The anisotropy factor g for the scattering medium was set to 0. A pixellated 2D contact detector with 20×20 pixels acquired the output image.

Our first objective was to investigate the performance of the proposed Neumann-series method in modeling photon transport. For this purpose, output images and fluence fields were generated using the proposed method. These were compared to the images and the fields obtained with the MCX Monte-Carlo (MC) technique [68]. For the MC study, the medium was discretized into 40 × 40 × 40 voxels. As in previous studies, the output using the MC technique was considered as the gold standard [54, 55]. The difference between the output obtained using the MC and the proposed Neumann-series technique over the entire image was quantified using the normalized root mean square error (RMSE) metric. Let M denote the number of pixels in the acquired image. In our setup, $M = 400$. Let $g_{m,RTE}$ and $g_{m,MC}$ denote the measurements acquired by the m^{th} detector pixel. Then the normalized RMSE was defined as follows:

$$RMSE = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(\frac{g_{m,RTE} - g_{m,MC}}{g_{m,MC}}^{2}\right)}. \tag{3.35}$$

We also examined the effect on accuracy of modeling photon transport when the reflection of photons at the boundary was not accounted. For this purpose, the output images were obtained using an existing Neumann-series method that assumed that the photons are completely absorbed at the boundary [54]. The RMSE for these output images were compared to those obtained using the proposed Neumann-series technique, thus quantifying the improvement in accuracy.

The second objective was to examine the effect on the accuracy of the estimated optical coefficients when the reflection of photons at the boundary was modeled. For this purpose, first the image data for the DOI setup in Fig. 3-4 was generated using the MC technique. Next, the absorption coefficient of the medium was estimated from this generated image using two reconstruction methods; the proposed Neumann-series-based reconstruction method and a modified Neumann-series-based reconstruction method that did not contain the reflection operator [54]. The scattering coefficient was assumed to be known. For this setup, the spatial basis function in equation (3.21) was simply the entire support of the object so that $N = 1$ and the objective was to only estimate a single value μ_a. The RMSE of the estimated absorption coefficient using the two reconstruction methods were compared.

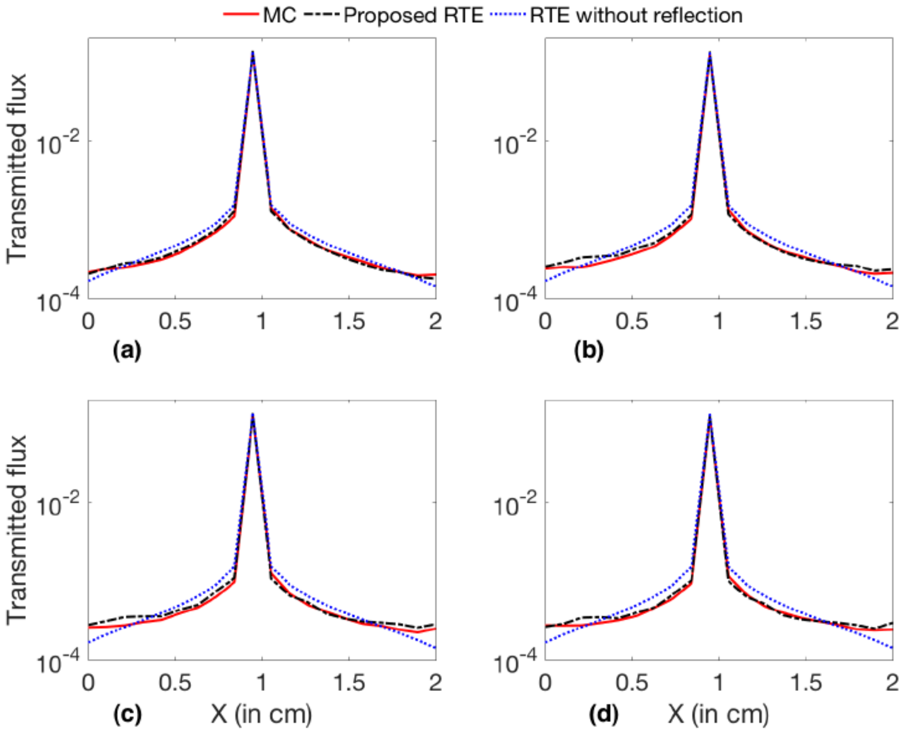


Figure 3-5. Comparison of the linear profiles of the images obtained with the MC and proposed Neumann-series RTE formalism, when the refractive index of the scattering medium is (a) 1.2 (b) 1.3 (c) 1.4 and (d) 1.5.

3.3 Results

Results from the experiments investigating the accuracy of photon transport are presented in Figs. 3-5 and 3-6. In Figs. 3-5a-d, the linear profile along the center of the output image obtained using the proposed Neumann series approach and using the MC technique are plotted for different amounts of refractive index mismatch. It is observed that the linear profiles overlap, demonstrating the accuracy of the proposed technique. The RMSE between the output images obtained using the MC and proposed Neumann-series RTE method is shown in Table 3-I. The corresponding RMSE using the existing Neumann-series technique that does not model the reflection of photons is also shown. It is observed that the RMSE values were lower using the proposed Neumann series technique with the average RMSE reduced by 38%.

Table 3-I. RMSE in the output image obtained using the Neumann-series RTE method with and without the reflection boundary conditions.

Refractive index n	RMSE for output image	
	Proposed Neumann-series RTE	Neumann-series RTE without reflection boundary conditions
1.1	14.2%	14.6%
1.2	9.5%	21.8%
1.3	14.5%	26.1%
1.4	15.6%	28.7%
1.5	18.1%	31.0%

The fluence on a 2D plane defined at $y = 0.9$ cm using the MC and proposed Neumann-series RTE method is shown in Fig. 3-6. In the third column of Fig. 3-6, the fluence contours for the Neumann-series RTE and the MC method are overlaid on top of each other, facilitating a visual comparison. It is observed that for different values of refractive index mismatch, the contour fields using the MC and RTE methods lie approximately on top of each other. Results from the experiments investigating the reconstruction method are presented in Fig. 3-7. It is observed that using the proposed reconstruction algorithm, the estimated absorption coefficient was close to the true absorption coefficient of 0.01 cm^{-1} for all values of refractive index, with an average normalized RMSE of 25.1%. However, on using the modified reconstruction algorithm that did not contain the reflection operator, the estimated absorption coefficient had a large error for all values of refractive index, with an average RMSE of 156.5%. On an average, we found that the mean RMSE of the estimated absorption coefficient reduced by 84%.

3.4 Discussions

The first goal of this work was to propose a Neumann-series RTE that modeled the reflection of photons at the boundary surface. The results in Fig. 3-5 and 3-6 demonstrate that the proposed Neumann-series method provided an accurate modeling of the photon transport for the considered experimental setup. Further, results in Table 3-I show that the proposed technique yields improved accuracy in modeling photon transport in comparison to when the reflection of photons at the boundary is not modeled. Thus, these results demonstrate that accounting for the reflection of photons at the boundary is essential to model photon transport accurately.

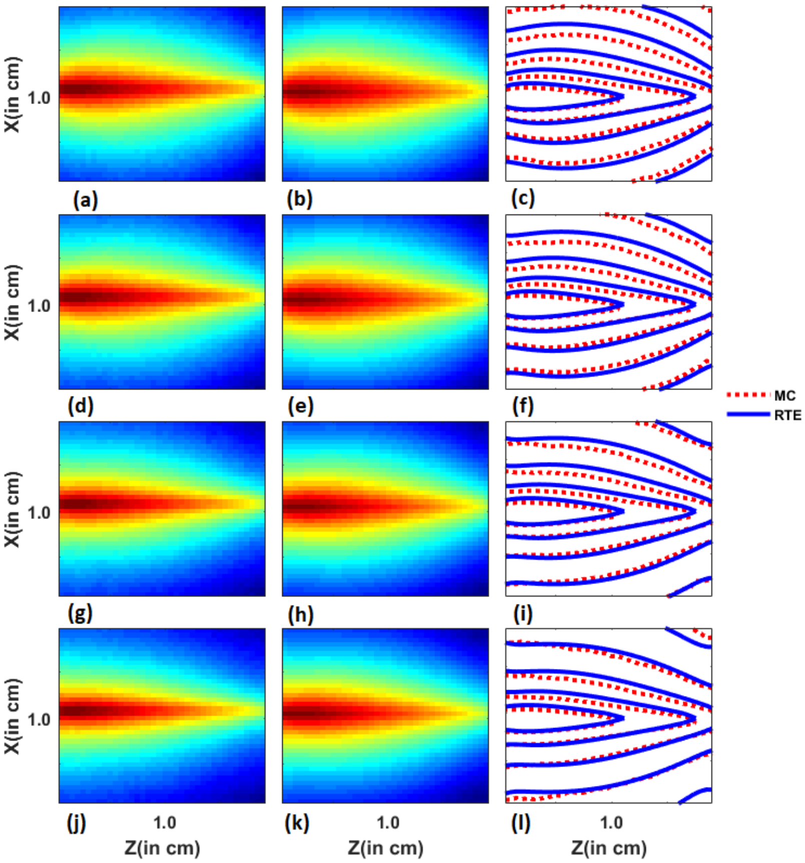
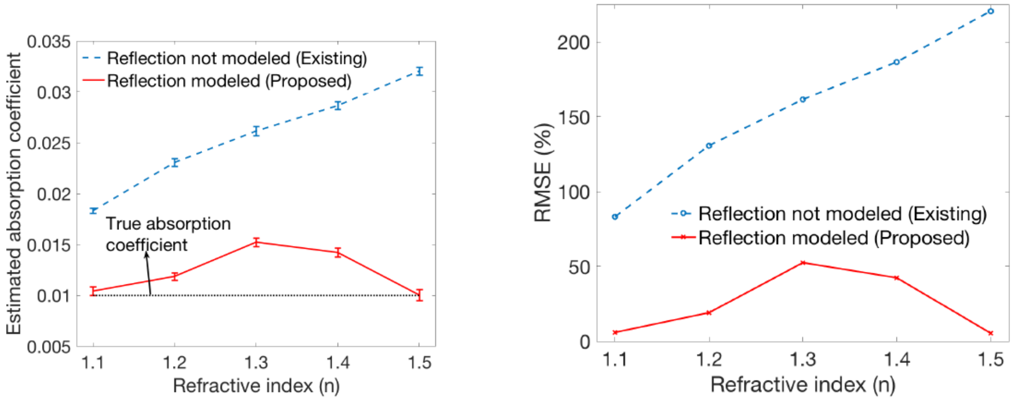


Figure 3-6. Comparison of the fluence calculated using the MC and proposed Neumann-series RTE-based methods. The refractive index of the medium is $n = 1.2$ (first row), $n = 1.3$ (second row), $n = 1.4$ (third row), and $n = 1.5$ (fourth row). In the first column (a, d, g, j), the logarithm of the amplitude of the fluence obtained with the MC method is plotted. In the second column (b, e, h, k), the logarithm of the amplitude of the fluence obtained with Neumann-series RTE is plotted. In the third column (c, f, i, l), the contour of the fluence obtained with the MC and Neumann-series RTE is plotted at 3 dB spacing.

The second goal of this work was to investigate the effect of modeling reflection at the boundary on the accuracy of the estimated optical coefficients. The results in Fig. 3-7 show that the optical coefficients estimated when the reflection was modeled



(a) (b)

Figure 3-7. (a) Plot of estimated absorption coefficient as a function of the refractive index using the proposed reconstruction method and another method that does not model photon reflection. Error bars are also shown. (b) Plot of the RMSE between the true and estimated absorption coefficient as a function of the refractive index using the two reconstruction methods

were substantially more accurate than when the reflection at the boundary was not modeled. These results demonstrate the importance of modeling the reflection of photons at the boundary of the tissue and the external medium.

To implement the proposed Neumann-series method, we considered only three additional terms in comparison to the original Neumann series. In each of these terms, the reflection operator was present once, under the assumption that terms that contain two or more instances of the reflection operator would not have a significant contribution to the output. This is because if the reflection coefficient R is not very high such that $R_2 \ll 1$, terms that contain the reflection operator twice could be ignored. For example, for the experimental setup in this manuscript, the plot of the reflection coefficient as a function of the incident angle is plotted in Fig. 3-8. We observe that the value of R is close to 0 for a majority of the angles of incidence. Another reason for choosing these three terms was that computing the other terms containing the reflection operator would require implementing the reflection operator in spherical harmonic basis. This is complicated because the reflection operator is highly directional. The results in Fig. 3-5 show that the output using the MC and RTE

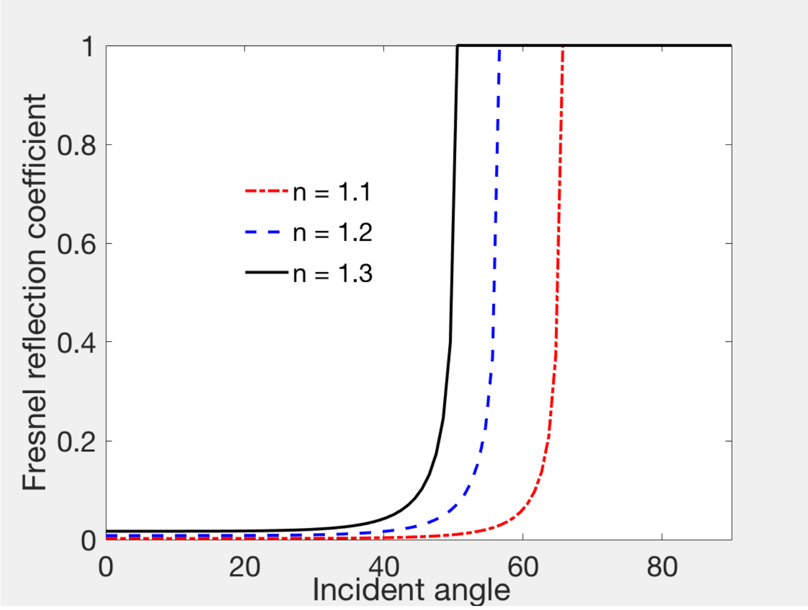


Figure 3-8. The Fresnel reflection coefficient values as a function of the incident angle for various values of refractive index n, when the surrounding media has a refractive index of 1.

approaches match well when only these three additional terms are used. Because the refractive index mismatch in the simulation study setup are representative of actual imaging setups, these results provide strong evidence for using only three additional terms in the Neumann-series RTE.

The Neumann-series method can handle different phase functions in a way similar to the other more conventional integro-differential methods to solve the RTE. This is because the scattering operator is in an integral form in both the Neumann series and the integro-differential form of the RTE. Also, the Neumann-series method method can handle complex illumination patterns. Note that the source term is modeled by the $\Xi(\boldsymbol{r}, \hat{\boldsymbol{s}})$ term in the Neumann-series RTE. The DOI setup used in this work consisted of a collimated monochromatic laser source, so that the source term was defined as in equation 3.13. For a different illumination pattern, this term would need to be modified accordingly. Finally, while the setup considered in this manuscript assumed a time-independent source, the Neumann-series RTE is applicable to time-dependent sources. For this purpose, note that the RTE can be represented in the frequency domain for each frequency component ν by replacing μ_{tot} in equation (3.1) by $\mu_{tot} + \frac{i\nu}{c_m}$, where i denotes the imaginary unit. Thus, the Neumann-series RTE can be solved for each frequency component by simply replacing μ_{tot} by $\mu_{tot} + \frac{i\nu}{c_m}$ [54]. Consequently, the proposed Neumann-series method could be used to model photon propagation in a wide range of imaging setups. We do note that for the different imaging setups, the number of terms required in the Neumann-series might be different and not known a priori. For this purpose, a criterion has been developed that enables assessing whether the Neumann-series RTE output has converged [54]. Using this criterion allows an adaptive determination of whether the Neumann series has converged or whether more terms need to be computed to obtain an accurate output.

For the experimental setup considered in this manuscript, the DA-based methods are inaccurate [50–54]. Our results demonstrate the accuracy and resultant advantage of the proposed Neumann series method in modeling photon transport in this setup. However, when the scattering coefficients are high, or the medium has a large geometry, the Neumann series method has large memory and computational requirements. Previously we have observed that the method is accurate and practical when $\mu_s L$ is around 5 [54], where L denotes the length of the medium. This is a limitation of the proposed method. In these cases, the DA-based methods yield accurate results. Thus, the proposed method could be integrated with the DA methods to yield hybrid approaches, similar to the RTE-DA method [187] and MC-DA method [188] proposed previously. Further, rapid advances in the high-performance computing hardware technology are being observed. For example, a NVIDIA graphics processing units (GPU)-based implementation of the RTE was proposed with the NVIDIA C2050 GPUs in 2012 [55]. The current generation of GPUs, namely the NVIDIA Volta, have about fifteen-times the processing power and five-times the memory of the C2050. Since the challenges with the Neumann-series method are mainly computational, we anticipate that these advances in high-performance computing will alleviate these challenges in the near future.

The implementation of the Neumann series approach with boundary conditions enables an experimental validation of this technique. This experimental validation is an important direction of future research. Another important frontier is the implementation of this method for heterogeneous medium. In this context, a Neumann-series method for a heterogeneous medium has been developed, but assumes vacuum boundary conditions [55]. Using the approaches outlined in this manuscript, the Neumann-series method for heterogeneous medium could be extended to model reflection of photons at the boundary of the tissue. Further, in several bio-photonics imaging applications, reflection can occur between different tissues. For example, in transcranial imaging, refractive index changes occur between skull and cerebrospinal fluid in the brain. Improving the Neumann-series approach to model photon propagation in such media is another important area of future study. Solving a set of coupled equations using a treatment similar to Lehtikangas et al. [189] offers a mechanism for this purpose.

The reconstruction method proposed in this manuscript was used to estimate the absorption coefficient for a medium that had uniform optical coefficients. This was because our main purpose was to investigate the effect of modeling boundary conditions on the accuracy of estimating the optical coefficients. However, the proposed reconstruction method is general and an important direction of research is to refine the method for estimating the scattering and attenuation coefficients for a non-uniform medium.

In this manuscript, we have focused on the deterministic approaches to solve the RTE. Another widely prevalent set of techniques for modeling photon transport are the stochastic MC-based techniques. The output of MC-based techniques is considered as the gold standard in several computational validation studies of the RTE, including in this manuscript. However, for image reconstruction, the use of stochastic MC-based techniques can lead to noisy estimates of the derivate or the need for simulating a large number of photons, which can be computationally expensive. In this context, several techniques have been proposed to accelerate the MC method [65]. However, using MC method to reconstruct the scattering coefficients is still challenging [190–192]. In contrast, deterministic approaches such as the Neumann-series RTE yield an analytical expression for the gradient that is not affected by noise. Given the trade-offs between the stochastic and deterministic RTE approaches, research on integrating these approaches for improved image reconstruction is an important frontier.

3.5 Conclusions

This work has proposed and investigated a Neumann-series-based RTE for improved modeling of photon propagation in tissue. The method models photon reflection at the interface of tissue and external medium using Fresnel's equations. Further, the method was used to develop an algorithm to reconstruct the absorption and scattering coefficients of a scattering medium. Computational studies demonstrated that the method yielded more accurate modeling of photon transport for a 3D diffuse optical imaging (DOI) system in comparison to when the photon reflection was not modeled. In addition, the method yielded substantially more accurate estimates of the absorption coefficients for the 3D DOI system. The results demonstrate the importance of accounting for the reflection of photons at boundary when modeling photon propagation.

3.A Derivation of the $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi$ term

Substituting the expression from equation (3.13) into equation (3.5) yields

$$[\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r},\hat{\boldsymbol{s}}) = \frac{\alpha}{c_m} h(x,y)\delta(\hat{\boldsymbol{s}}-\hat{\boldsymbol{z}}) \exp\left[-\int_0^z \mu_{tot}(\boldsymbol{r}-\hat{\boldsymbol{z}}\lambda)\right]. \tag{3.36}$$

The photons described by the $\boldsymbol{\mathcal{X}}\Xi$ term are incident on the boundary defined by the plane $z = L$. We define this infinitesimally thin boundary by the expression $\delta(z - L)$. Applying the reflection operator, as defined by equation (3.8), on the distribution function within this boundary yields

$$[\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \int_{\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{s}}'>0} d\Omega' \alpha h(x,y)\delta(z-L)R(\hat{\boldsymbol{s}}')\delta[\hat{\boldsymbol{s}} - \hat{\boldsymbol{s}}' + 2(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{s}}')\hat{\boldsymbol{n}}]\delta(\hat{\boldsymbol{s}}' - \hat{\boldsymbol{z}})\times \\ \exp\left[-\int_0^L d\lambda \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{z}}\lambda)\right]. \tag{3.37}$$

Using the sifting property of the delta function and the fact that the normal to the boundary surface $\hat{\boldsymbol{n}}$ is the direction vector $\hat{\boldsymbol{z}}$ yields

$$[\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \alpha h(x,y)R(\hat{\boldsymbol{z}})\delta(\hat{\boldsymbol{s}} + \hat{\boldsymbol{z}})\delta(z-L)\exp\left[-\int_0^L d\lambda \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{z}}\lambda)\right]. \tag{3.38}$$

The above expression has a simple physical interpretation, namely that the photons are reflected back into the medium along the negative $\hat{\boldsymbol{z}}$ direction, as would be expected. Finally, applying the attenuation operator in equation (3.5) yields

$$[\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \frac{\alpha R(\hat{\boldsymbol{z}})}{c_m} h(x,y)\delta(\hat{\boldsymbol{s}} + \hat{\boldsymbol{z}})\exp\left[-\int_0^L d\lambda' \mu_{tot}(x, y, L - \lambda') - \\ \int_0^{L-z} d\lambda' \mu_{tot}(\boldsymbol{r} + \hat{\boldsymbol{z}}\lambda')\right]. \tag{3.39}$$

3.B Derivation of the $\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi$ term

In [54], it was shown

$$[\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \alpha \mu_s(\boldsymbol{r}) h(x,y) p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{z}}) \exp\left[-\int_0^z d\lambda' \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{z}}\lambda')\right]. \tag{3.40}$$

With equation (3.5), we obtain

$$[\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \frac{\alpha}{c_m}\int_0^\infty d\lambda \mu_s(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda) h(x - \hat{\boldsymbol{s}}_x\lambda, y - \hat{\boldsymbol{s}}_y\lambda) p(\hat{\boldsymbol{s}}, \hat{\boldsymbol{z}})\times \\ \exp\left[-\int_0^{z - \hat{\boldsymbol{s}}_z\lambda} d\lambda' \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda - \hat{\boldsymbol{z}}\lambda') - \int_0^\lambda d\lambda' \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda')\right]. \tag{3.41}$$

Next, the reflection operator must be applied. This operator acts only on locations within a thin boundary region, denoted by $\sum_i \delta(\boldsymbol{r}\cdot\hat{\boldsymbol{n}}_i - p_i)$. Applying the reflection operator, as defined by equation (3.8), on $[\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r},\hat{\boldsymbol{s}})$ with this definition for the boundary yields

$$[\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r},\hat{\boldsymbol{s}}') = \alpha\sum_i \int_{\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{s}}>0} d\Omega \int_0^\infty d\lambda \delta(\boldsymbol{r}\cdot\hat{\boldsymbol{n}}_i - p_i) R(\hat{\boldsymbol{s}})\delta(\hat{\boldsymbol{s}}' - \hat{\boldsymbol{s}} + 2(\hat{\boldsymbol{n}}_i\cdot\hat{\boldsymbol{s}})\hat{\boldsymbol{n}}_i)\times \\ \mu_s(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda)h(x-\hat{\boldsymbol{s}}_x\lambda, y-\hat{\boldsymbol{s}}_y\lambda)p(\hat{\boldsymbol{s}},\hat{\boldsymbol{z}})\times \\ \exp\left[-\int_0^{z-\hat{\boldsymbol{s}}_z\lambda} d\lambda'\mu_{tot}(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda-\hat{\boldsymbol{z}}\lambda') - \int_0^\lambda d\lambda'\mu_{tot}(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda')\right]. \tag{3.42}$$

To simplify the above equation, replace $\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda$ by $\boldsymbol{r}'$, so that $\hat{\boldsymbol{s}} = \frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|}$, $\lambda = |\boldsymbol{r}-\boldsymbol{r}'|$ and $d^3r' = \lambda^2 d\lambda d\Omega$. For notational simplicity, define

$$\hat{\boldsymbol{s}}_{10} = \frac{\boldsymbol{r}-\boldsymbol{r}'}{|\boldsymbol{r}-\boldsymbol{r}'|}, \tag{3.43}$$

where we recognize the fact that $\hat{\boldsymbol{s}}_{10}$ is not a constant, but a function of r and r'. This yields

$$[\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r},\hat{\boldsymbol{s}}') = \alpha\sum_i \int d^3r' \frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|^2} R(\hat{\boldsymbol{s}}_{10})\delta(\hat{\boldsymbol{s}}' - \hat{\boldsymbol{s}}_{10} + 2(\hat{\boldsymbol{n}}_i\cdot\hat{\boldsymbol{s}}_{10})\hat{\boldsymbol{n}}_i)\times \\ \mu_s(\boldsymbol{r}')h(x',y')p(\hat{\boldsymbol{s}}_{10},\hat{\boldsymbol{z}})\delta(\boldsymbol{r}\cdot\hat{\boldsymbol{n}}_i - p_i)\times \\ \exp\left[-\int_0^{z'} d\lambda'\mu_{tot}(\boldsymbol{r}'-\hat{\boldsymbol{z}}\lambda') - \int_0^\lambda d\lambda'\mu_{tot}(\boldsymbol{r}-\hat{\boldsymbol{s}}_{10}\lambda')\right]. \tag{3.44}$$

With attenuation operator defined in (3.5), we have

$$[\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}\boldsymbol{\mathcal{X}}\Xi](\boldsymbol{r},\hat{\boldsymbol{s}}') = \alpha\sum_i \int_0^\infty d\lambda'' \int d^3r' \frac{1}{|\boldsymbol{r}-\hat{\boldsymbol{s}}'\lambda''-\boldsymbol{r}'|^2}\mu_s(\boldsymbol{r}')R(\hat{\boldsymbol{s}}_{10})\times \\ \delta(\hat{\boldsymbol{s}}' - \hat{\boldsymbol{s}}_{10} + 2(\hat{\boldsymbol{n}}_i\cdot\hat{\boldsymbol{s}}_{10})\hat{\boldsymbol{n}}_i)h(x',y')p(\hat{\boldsymbol{s}}_{10},\hat{\boldsymbol{z}})\delta(\boldsymbol{r}\cdot\hat{\boldsymbol{n}}_i - \hat{\boldsymbol{s}}'\cdot\hat{\boldsymbol{n}}_i\lambda'' - p_i)\times \\ \exp\left[-\int_0^{z'} d\lambda'\mu_{tot}(\boldsymbol{r}'-\hat{\boldsymbol{z}}\lambda') - \int_0^\lambda d\lambda'\mu_{tot}(\boldsymbol{r}-\hat{\boldsymbol{s}}'\lambda''-\hat{\boldsymbol{s}}_{10}\lambda')\right]. \tag{3.45}$$

Using the scaling property of the delta function and we have

$$\delta(\boldsymbol{r}\cdot\hat{\boldsymbol{n}}_i - \hat{\boldsymbol{s}}'\cdot\hat{\boldsymbol{n}}_i\lambda'' - p_i) = \frac{1}{\hat{\boldsymbol{s}}'\cdot\hat{\boldsymbol{n}}_i}\delta\left(\frac{\boldsymbol{r}\cdot\hat{\boldsymbol{n}}_i - p_i}{\hat{\boldsymbol{s}}'\cdot\hat{\boldsymbol{n}}_i} - \lambda''\right). \tag{3.46}$$

For notational simplicity, define

$$\lambda_{21i} = \frac{\boldsymbol{r}\cdot\hat{\boldsymbol{n}}_i - p_i}{\hat{\boldsymbol{s}}'\cdot\hat{\boldsymbol{n}}_i}, \tag{3.47}$$

we recognize that λ_{21i} is a function of $\boldsymbol{r}$ and $\hat{\boldsymbol{s}}$. Substituting (3.46) in (3.45) and using the above definition for λ_{21i} yields

$$[\boldsymbol{\mathcal{XRXKX}}\Xi](\boldsymbol{r}, \hat{\boldsymbol{s}}') = \alpha \sum_i \int d^3r' \frac{1}{|\boldsymbol{r} - \hat{\boldsymbol{s}}'\lambda_{21i} - \boldsymbol{r}'|^2} \frac{1}{\hat{\boldsymbol{s}}' \cdot \hat{\boldsymbol{n}}_i} \mu_s(\boldsymbol{r}') R(\hat{\boldsymbol{s}}_{10}) \times \\ \delta(\hat{\boldsymbol{s}}' - \hat{\boldsymbol{s}}_{10} + 2(\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{s}}_{10})\hat{\boldsymbol{n}}_i) h(x', y') p(\hat{\boldsymbol{s}}_{10}, \hat{\boldsymbol{z}}) \times \\ \exp\left[-\int_0^{z'} d\lambda' \mu_{tot}(\boldsymbol{r}' - \hat{\boldsymbol{z}}\lambda') - \int_0^{\lambda} d\lambda' \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{s}}'\lambda_{21i} - \hat{\boldsymbol{s}}_{10}\lambda')\right]. \tag{3.48}$$

Substituting the above expression in equation (3.20) yields the following expression for the transmitted flux detected by the mth pixel due to the $\boldsymbol{\mathcal{XRXKX}}\Xi$ term:

$$\Phi_m = c_m \Delta x \Delta y \sum_i T(\hat{\boldsymbol{s}}_{21i})(\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{s}}_{21i}) \int d^3r' \frac{1}{|\boldsymbol{r}_m - \hat{\boldsymbol{s}}_{21i}\lambda_{21i} - \boldsymbol{r}'|^2} \frac{1}{\hat{\boldsymbol{s}} \cdot \hat{\boldsymbol{n}}_i} \mu_s(\boldsymbol{r}') R(\hat{\boldsymbol{s}}_{10}) \times \\ h(x', y') p(\hat{\boldsymbol{s}}_{10}, \hat{\boldsymbol{z}}) \exp\left[-\int_0^{z'} d\lambda' \mu_{tot}(\boldsymbol{r}' - \hat{\boldsymbol{z}}\lambda') - \int_0^{\lambda} d\lambda' \mu_{tot}(\boldsymbol{r}_m - \hat{\boldsymbol{s}}_{21i}\lambda_{21i} - \hat{\boldsymbol{s}}_{10}\lambda')\right], \tag{3.49}$$

where we have used the sifting property of the delta function and where

$$\hat{\boldsymbol{s}}_{21i} = \hat{\boldsymbol{s}}_{10} - 2(\hat{\boldsymbol{n}}_i \cdot \hat{\boldsymbol{s}}_{10})\hat{\boldsymbol{n}}_i. \tag{3.50}$$

3.C Derivation of the $\frac{\partial \boldsymbol{\mathcal{X}}}{\partial \mu_n}\boldsymbol{\mathcal{R}}$ term

Applying the attenuation operator in equation (3.5) on the expression for the $\boldsymbol{\mathcal{R}}w$ term in (3.8), we have

$$[\boldsymbol{\mathcal{XR}}w](\boldsymbol{r}, \hat{\boldsymbol{s}}) = \frac{1}{c_m} \int_0^{\infty} d\lambda [\boldsymbol{\mathcal{R}}w](\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda, \hat{\boldsymbol{s}}) \exp\left[-\int_0^{\lambda} d\lambda' \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda')\right]. \tag{3.51}$$

Taking the derivative of the above expression with respect to μ_n yields

$$\frac{\partial \boldsymbol{\mathcal{X}}}{\partial \mu_n} \boldsymbol{\mathcal{R}}w(\boldsymbol{r}, \hat{\boldsymbol{s}}) = -\frac{1}{c_m} \int_0^{\infty} d\lambda [\boldsymbol{\mathcal{R}}w](\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda, \hat{\boldsymbol{s}}) \exp\left[-\int_0^{\lambda} d\lambda' \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda')\right] \left[\int_0^{\lambda} d\lambda'' \phi_n(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda'')\right]. \tag{3.52}$$

By rearranging terms followed by a change of integration order, which is similar as described in [186], we have

$$\frac{\partial \boldsymbol{\mathcal{X}}}{\partial \mu_n} \boldsymbol{\mathcal{R}}w(\boldsymbol{r}, \hat{\boldsymbol{s}}) = -\frac{1}{c_m} \int_0^{\infty} d\lambda'' \phi_n(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda'') \int_{\lambda''}^{\infty} d\lambda [\boldsymbol{\mathcal{R}}w](\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda, \hat{\boldsymbol{s}}) \exp\left[-\int_0^{\lambda} d\lambda' \mu_{tot}(\boldsymbol{r} - \hat{\boldsymbol{s}}\lambda')\right]. \tag{3.53}$$

Separating the exponential integral into two parts, we obtain

$$\begin{aligned}\frac{\partial \boldsymbol{\mathcal{X}}}{\partial \mu_n}\boldsymbol{\mathcal{R}}w(\boldsymbol{r},\hat{\boldsymbol{s}}) = -\frac{1}{c_m}\int_0^\infty d\lambda''\phi_n(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda'')\exp\left[-\int_0^{\lambda''} d\lambda'\mu_{tot}(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda')\right]\times \\ \int_{\lambda''}^\infty d\lambda[\boldsymbol{\mathcal{R}}w](\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda,\hat{\boldsymbol{s}})\exp\left[-\int_{\lambda''}^{\lambda} d\lambda'\mu_{tot}(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda'\right].\end{aligned} \tag{3.54}$$

Note that

$$\int_{\lambda''}^\infty d\lambda[\boldsymbol{\mathcal{R}}w](\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda,\hat{\boldsymbol{s}})\exp\left[-\int_{\lambda''}^{\lambda} d\lambda'\mu_{tot}(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda')\right] = c_m\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}w(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda'',\hat{\boldsymbol{s}}). \tag{3.55}$$

Thus

$$\begin{aligned}\frac{\partial \boldsymbol{\mathcal{X}}}{\partial \mu_n}\boldsymbol{\mathcal{R}}w(\boldsymbol{r},\hat{\boldsymbol{s}}) &= -\frac{1}{c_m}\int_0^\infty d\lambda''\phi_n(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda'')c_m\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}w(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda'',\hat{\boldsymbol{s}})\exp\left[-\int_0^{\lambda''} d\lambda'\mu_{tot}(\boldsymbol{r}-\hat{\boldsymbol{s}}\lambda')\right] \\ &= -c_m\boldsymbol{\mathcal{X}}(\phi_n\boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}w).\end{aligned} \tag{3.56}$$

Chapter 4

Sparsity-initialized maximum-likelihood expectation maximization for fluorescence molecular tomography reconstruction

In this chapter, we present a reconstruction method involving maximum-likelihood expectation maximization (MLEM) to model Poisson noise as applied to fluorescence molecular tomography (FMT). MLEM is initialized with the output from a sparse reconstruction-based approach, which performs truncated singular value decomposition-based preconditioning followed by fast iterative shrinkage-thresholding algorithm (FISTA) to enforce sparsity. The motivation for this approach is that sparsity information could be accounted for within the initialization, while MLEM would accurately model Poisson noise in the FMT system. Simulation experiments show the proposed method significantly improves images qualitatively and quantitatively. The method results in over 20 times faster convergence compared to uniformly initialized MLEM and improves robustness to noise compared to pure sparse reconstruction. We also theoretically justify the ability of the proposed approach to reduce noise in the background region compared to pure sparse reconstruction. Overall, these results provide strong evidence to model Poisson noise in FMT reconstruction and for application of the proposed reconstruction framework to FMT imaging.

4.1 Introduction

Fluorescence molecular tomography (FMT) is finding several applications in 3D visualization and quantification of the distribution of molecular target within biological tissue [74]. In particular, FMT has received substantial interest in small animal imaging for applications such as studying tumor physiology and for pharmaceutical research [8, 193]. In FMT imaging, fluorescence molecules are first injected into biological tissue. External illumination sources are used to excite the fluorescence molecules. The photons emitted by the excited fluorescence molecules are collected by detectors at the tissue surface. The objective in FMT is to use these surface measurements to reconstruct the 3D distribution of fluorescence molecules within the tissue.

The reconstruction problem in FMT is known to be highly ill-posed, and is sensitive to noise and modeling errors such as discretization [72, 194]. Over the past two decades, various reconstruction methods for FMT have been proposed [49]. Tikhonov regularization is a popular regularization applied to FMT reconstruction problem. The regularized problem can be solved iteratively with methods such as Newton method and algebraic reconstruction technique (ART) [49, 195]. However, such regularization tends to over-smooth the reconstructed images, leading to loss of localized features during reconstruction [196]. More recently, reconstruction methods that exploit sparsity of the fluorescence distribution have been studied [72, 83, 197, 198]. In these methods, ℓ_0 or ℓ_1 regularization on the fluorescence distribution is applied to enforce sparsity while performing the reconstruction. These regularization problems can be solved with methods such as greedy algorithms and iterative thresholding methods [199].

In FMT systems, often the detector system is charged-couple device (CCD)-based camera or photon multiplier tube (PMT). In these systems, the noise is described by a Poisson distribution [200–202]. For this noise distribution, MLEM-based reconstruction techniques have yielded reliable results, especially in nuclear medicine imaging [45, 203–206]. The MLEM technique has several advantages, such as accurately modeling the Poisson noise distribution in the acquired data, constraining the activity values to be non-negative without the need for a specific regularizer, and ensuring the conservation of the total number of photons across multiple iterations. In optical tomography, several studies have applied MLEM for reconstruction in bioluminescence tomography [207–209]. In [210], MLEM has also been applied for FMT reconstruction. However, the MLEM technique typically suffers from slow convergence for optical tomography modalities, with thousands of iterations and large amount of time per iteration being required [209, 211, 212]. This makes MLEM a time-consuming method and thus not very practical [207, 210]. As a result, MLEM has not been widely used for image reconstruction in optical tomography.

The performance of MLEM is influenced by different factors. An important factor being the initial estimate provided to the algorithm. Conventionally, MLEM starts with a uniform initial estimate, as we explain later. However, different initializations for MLEM yield different reconstruction results [45, 213]. In this work, we studied the use of sparse reconstruction to initialize the MLEM approach. The overall motivation for this approach is that the sparse reconstruction method would account for the sparsity of the fluorescence distribution, while the MLEM would accurately model the Poisson noise in the FMT system. However, this combined approach is also able to exploit several inherent advantages of these two techniques, as we describe below. Our method yields reliable and improved results in comparison to pure sparse reconstruction as well as uniformly initialized MLEM methods.

4.2 Methods

4.2.1 The forward model and reconstruction problem in FMT

As we have discussed in 2.1.3, the forward model in FMT is described by a pair of coupled equations. The first equation describes the propagation of excitation photons from source at location $\boldsymbol{r}_s$ to location $\boldsymbol{r}$ in the medium and the second one describes the propagation of emitted fluorescence photons from location $\boldsymbol{r}$ to detector at location $\boldsymbol{r}_d$, where $\boldsymbol{r}_s$, $\boldsymbol{r}$ and $\boldsymbol{r}_d$ are three-dimensional vectors. Using $\phi_{ex}(\boldsymbol{r})$ and $\phi_{em}(\boldsymbol{r})$ to represent excitation light field at $\boldsymbol{r}$ and emission light field at $\boldsymbol{r}_d$, we have

$$\phi_{ex}(\boldsymbol{r}) = \int_{\Omega} g_{ex}(\boldsymbol{r}_s, \boldsymbol{r}) s(\boldsymbol{r}_s) d\boldsymbol{r}_s, \tag{4.1}$$

and

$$\phi_{em}(\boldsymbol{r}_d) = \int_{\Omega} g_{em}(\boldsymbol{r}, \boldsymbol{r}_d) x(\boldsymbol{r}) \phi_{ex}(\boldsymbol{r}) d\boldsymbol{r}, \tag{4.2}$$

where $g_{ex}(\boldsymbol{r}_s, \boldsymbol{r})$ and $g_{em}(\boldsymbol{r}, \boldsymbol{r}_d)$ are Green's function of excitation light and emission light, respectively. $x(\boldsymbol{r})$ is the fluorescence yield at location $\boldsymbol{r}$, and Ω denotes object support. If we discretize Ω into N voxels, we obtain the linear matrix equation for the forward model:

$$\boldsymbol{\Phi} = \mathbf{G}\boldsymbol{x}, \tag{4.3}$$

where

$$\mathbf{G} = \begin{bmatrix} g^{1}_{em,1}\phi^{1}_{em,1} & \cdots & g^{1}_{em,N}\phi^{1}_{em,N} \\ \vdots & & \vdots \\ g^{N_d}_{em,1}\phi^{1}_{em,1} & \cdots & g^{N_d}_{em,N}\phi^{1}_{em,N} \\ g^{1}_{em,1}\phi^{2}_{em,1} & \cdots & g^{1}_{em,N}\phi^{2}_{em,N} \\ \vdots & & \vdots \\ g^{N_d}_{em,1}\phi^{N_s}_{em,1} & \cdots & g^{N_d}_{em,N}\phi^{N_s}_{em,N} \end{bmatrix}$$

is the sensitivity matrix of the system. Assuming we have N_s laser sources and N_d detectors, $\boldsymbol{\Phi}$ is an $M \times 1$ vector with $M = N_s \times N_d$ representing the total number of measurements. $\boldsymbol{x}$ is an $N \times 1$ vector representing unknown fluorescence yield. Due to the limited number of sources and detectors, typically $M < N$ in FMT.

Modeling the measurement noise denoted by the M-dimensional vector $\boldsymbol{n}$, equation (4.4) becomes

$$\boldsymbol{\Phi} = \mathbf{G}\boldsymbol{x} + \boldsymbol{n}, \tag{4.4}$$

In FMT, the data collected by the detectors is corrupted by Poisson noise [200]. The reconstruction problem in FMT is to reconstruct $\boldsymbol{x}$ given sensitivity matrix $\mathbf{G}$ and detector measurements $\boldsymbol{\Phi}$.

4.2.2 Modeling Poisson noise in the reconstruction

We have shown in 2.2.3.2 that the MLEM can be derived from Poisson likelihood function. Here we just rewrite the MLEM update as follows:

$$\hat{x}_n^{(k+1)} = \hat{x}_n^{(k)} \frac{1}{s_n} \sum_{m=1}^{M} \frac{\phi_m}{(\mathbf{G}\hat{\boldsymbol{x}}^{(k)})_m} G_{mn}, \tag{4.5}$$

where $s_n = \sum_{m=1}^{M} G_{mn}$.

The MLEM iteration starts from an initial estimate $\hat{\boldsymbol{x}}^{(0)}$, and and the results of this technique can be influenced by this initial estimate [45]. Typically, the initial estimate is uniform, where all the elements in $\hat{\boldsymbol{x}}^{(0)}$ are assumed to be a constant [214, 215]. However, with this estimate, MLEM updates all the voxels in every iteration, increasing the computational requirements. In equation (4.5), note that $\hat{x}_n^{(k)}$ will always be zero if $\hat{x}_n^{(0)} = 0$ due to the multiplicative nature of the technique. Thus, the zero elements can be excluded from $\hat{\boldsymbol{x}}^{(0)} = 0$ during the MLEM iteration. Matrix $\mathbf{G}$ used for MLEM iteration can be formulated with columns corresponding to non-zero elements in $\hat{\boldsymbol{x}}^{(0)}$. This reduces the size of matrices in the reconstruction problem and accelerates the computation speed. In this context, in many FMT applications, fluorescence molecules tend to concentrate in a small target region. Thus, if we could exploit this property, we could generate a sparse initial estimate, which allows us to accelerate the MLEM technique. Such a technique would inherently exploit the sparsity-based prior information in FMT as well as model the Poisson noise in FMT accurately. Inspired by this, we developed a sparse reconstruction method and used the output from this method as the initial estimate for MLEM. In the next section, we describe the method we used to obtain sparse initial estimate of MLEM.

4.2.3 Sparse reconstruction and preconditioning of sensitivity matrix

In order to obtain the sparse initial estimate for MLEM, the following minimization problem can be formulated based on equation (4.4):

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_0 \quad \text{such that} \quad \|\boldsymbol{\Phi} - \mathbf{G}\boldsymbol{x}\|_2 \leq \epsilon. \tag{4.6}$$

While directly solving this problem is computationally complex, equation (4.6) can be approximately solved with greedy algorithms or convex relaxation techniques [199]. The theory of compressed sensing (CS) provides the conditions under which such approximate solvers are valid. Further, approaches based on singular value decomposition (SVD) can be applied to the sensitivity matrix to improve sparse reconstruction in FMT [38, 69, 216, 217]. This technique is known as preconditioning of sensitivity matrix. Here, we follow truncated singular value decomposition (TSVD) described in [217] as the preconditioning method. First, expressing the matrix G in terms of its singular vectors and singular values using SVD, equation (4.4) becomes:

$$\boldsymbol{\Phi} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T\boldsymbol{x} + \boldsymbol{n}, \tag{4.7}$$

where $\boldsymbol{U}$ and $\boldsymbol{V}$ are $M \times M$ and $N \times N$ unitary matrices where the columns are left-singular vectors and right-singular vectors, respectively, and $\boldsymbol{\Sigma}$ is a diagonal matrix where the diagonal elements are the singular values. By multiplying both sides of equation (4.7) with $\boldsymbol{\Sigma}^{-1}\boldsymbol{U}^T$, we could potentially use $\boldsymbol{V}^T$ as the new sensitivity matrix. However, since the reconstruction problem in FMT is highly ill-posed, the inversion of small singular values contained in $\boldsymbol{\Sigma}$ will cause large noise amplification. To address this issue, we keep only the K largest singular values of matrix $\boldsymbol{\Sigma}$ and discard the rest, before performing the inversion of $\boldsymbol{\Sigma}$. The corresponding columns in $\boldsymbol{U}$ and $\boldsymbol{V}$ are also discarded. This process is referred to as truncation. Then (4.7) becomes

$$\boldsymbol{\Phi} = \boldsymbol{U}_t \boldsymbol{\Sigma}_t \boldsymbol{V}_t^T \boldsymbol{x} + \boldsymbol{n}, \tag{4.8}$$

where the size of $\boldsymbol{U}_t$, $\boldsymbol{\Sigma}_t$ and $\boldsymbol{V}_t$ are $M \times K$, $K \times K$ and $N \times K$, respectively. Since small singular values are discarded, usually $K < M$. Applying $\boldsymbol{M} = \boldsymbol{\Sigma}_t^{-1} \boldsymbol{U}_t^T$ to both sides of equation (4.8) yields

$$\boldsymbol{M\Phi} = \boldsymbol{V}_t^T \boldsymbol{x} + \boldsymbol{Mn}. \tag{4.9}$$

Denoting $\boldsymbol{y} = \boldsymbol{M\Phi}$, $\boldsymbol{A} = \boldsymbol{V}_t^T$ and $\boldsymbol{n}' = \boldsymbol{Mn}$, equation (4.9) can be written as

$$\boldsymbol{y} = \boldsymbol{Ax} + \boldsymbol{n}'. \tag{4.10}$$

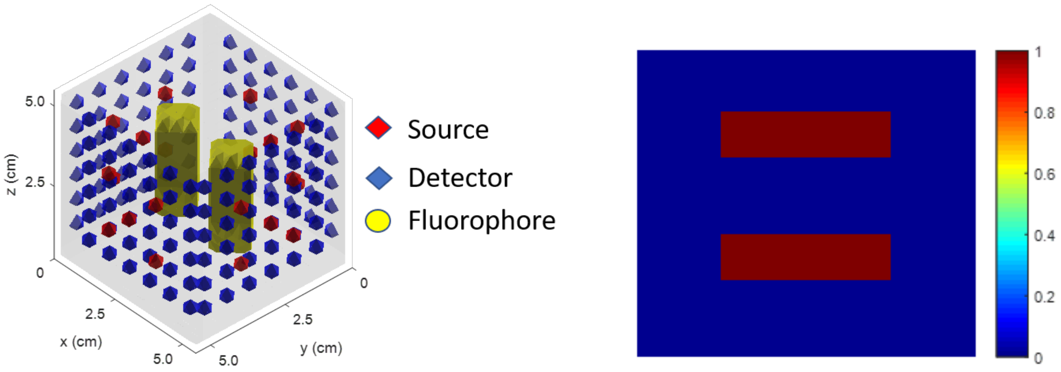
We now solve equation (4.10) as a sparse reconstruction problem. More specifically, we implemented convex relaxation technique in this work. Our objective is to minimize the ℓ_1 norm of the vector $\boldsymbol{x}$. Thus the sparse reconstruction problem is posed as

$$\min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \quad \text{such that} \quad \|\boldsymbol{y} - \boldsymbol{Ax}\|_2 \leq \epsilon. \tag{4.11}$$

We applied the fast iterative shrinkage-thresholding algorithm (FISTA) for solving the minimization problem in equation (4.11) [218]. The output with this method is then input to the MLEM technique as the initial estimate. Note that results from sparse reconstruction might contain negative elements. As we explained previously, MLEM constrains the activity values to be non-negative. To enable this, the negative elements in $\hat{\boldsymbol{x}}^{(0)}$ are set to zero.

4.2.4 Experiments

To validate the proposed method, different simulation experiments were conducted. Three different reconstruction methods were implemented for comparison, namely, (a) a pure sparsity-based reconstruction method that used TSVD in conjunction with

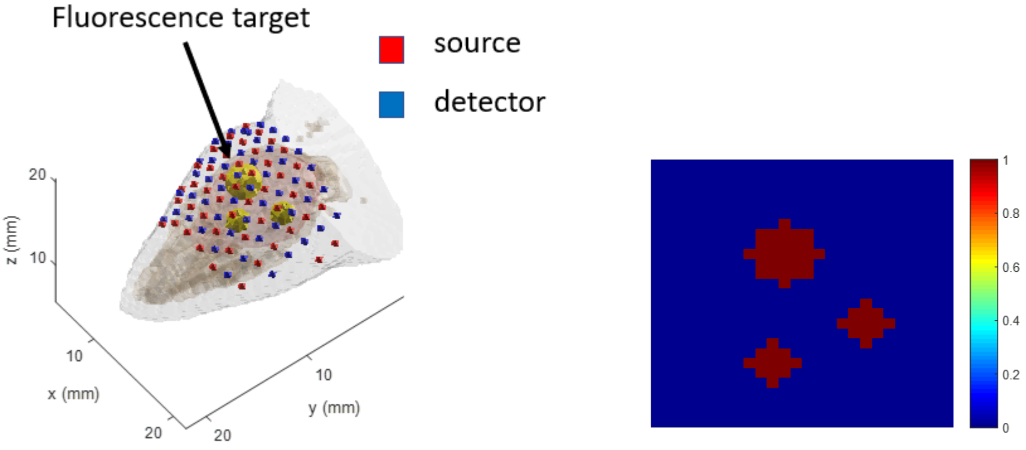


(a) **(b)**

Figure 4-1. (a) The experimental setup of cube phantom. (b) Cross section at $y = 2.5$ cm of the simulated phantom.

FISTA, (b) the MLEM method with uniform initial estimate of the image (more specifically, the initial activity values in all the voxels was set to unity) and (c) the MLEM method with an initialization that was obtained using the method described in (a). We will refer to these methods as pure sparsity-based reconstruction method, uniformly initialized MLEM and sparsity-initialized MLEM, respectively.

In the first set of experiments, a $5 \times 5 \times 5$ cm^3 cubic phantom was considered, as shown in Fig. 4-1(a). The phantom was discretized into $20 \times 20 \times 20$ voxels. The absorption coefficient of the phantom was set to $\mu_a = 0.05$ cm^{-1} and the reduced scattering coefficient was set to $\mu_s' = 10$ cm^{-1}. 20 sources and 144 detectors were placed on the side surfaces. This configuration generated 2880 measurements. Two cylindric fluorescence bars with radius of 0.375 cm and length of 2.5 cm each were inserted into the phantom. The fluorescence intensity in these bars was set to unity. The cross section of the phantom at $y = 2.5$ cm is shown in Fig. 4-1(b). The Green's function in the forward model of FMT was computed using Monte Carlo method, where a large number of photons were simulated to generate approximately noiseless measurements[68]. The measurements were then scaled to different levels and corresponding Poisson noise was applied using a Poisson distributed pseudo random number generator. This yielded detector measurements with different signal-to-noise



(a)

(b)

Figure 4-2. (a) The experimental setup of digital mouse phantom. (b) Cross section of digital mouse phantom at $z = 16\text{mm}$.

ratio (SNR) values.

Table 4-I. Optical properties of digital mouse phantom[219]

Tissue type	Brain	Skull	Skin
$\mu_s'(\text{ cm}^{-1})$	12.5	10.0	8.0
$\mu_a(\text{ cm}^{-1})$	0.178	0.101	0.159

To study the effect of MLEM iteration number on reconstruction performance, 1000 iterations were performed for MLEM with different initializations with the SNR initially set to 18dB, and the truncation number K set to 760. The region of interest (ROI) corresponded to the region occupied by the fluorescence bars. The rest of the region was defined as background. For quantitative study, different figures of merit were computed. Specifically, we computed absolute bias in the estimated uptake in the ROI and the background, spatial variance within the pixels in the ROI and the background, and the root mean square error (RMSE) for the entire image. The mean of the fluorescence uptake within the ROI, denoted by θ_{ROI}, is defined as

$$\theta_{\text{ROI}} = \frac{1}{N_R}\sum_{r=1}^{N_R} x_r, \tag{4.12}$$

where r denotes the r^{th} voxel in the ROI, and N_R is the number of voxels in the ROI.

Similarly, the background mean, denoted by θ_{B}, is defined as

$$\theta_{\mathrm{B}} = \frac{1}{N_B}\sum_{b=1}^{N_B} x_b, \tag{4.13}$$

where b denote the b^{th} voxel in the background region, and N_B is the number of voxels in the background. Then the ROI absolute bias, denoted by b_{ROI}, was computed as:

$$b_{\mathrm{ROI}} = \frac{1}{R}\sum_{k=1}^{R} |\theta_{\mathrm{ROI},k} - \theta_{\mathrm{ROI},k}^{true}|, \tag{4.14}$$

where k denotes the k^{th} noise realization, $\theta_{\mathrm{ROI},k}^{true}$ denotes the true uptake in the k^{th} voxel in the ROI, and R is the total number of noise realizations. The background absolute bias, denoted by b_{B}, was computed as:

$$b_{\mathrm{B}} = \frac{1}{R}\sum_{k=1}^{R} |\theta_{\mathrm{B},k} - \theta_{\mathrm{B},k}^{true}|, \tag{4.15}$$

where $\theta_{\mathrm{B},k}^{true}$ denotes the true uptake in the k^{th} voxel in the background. We also computed the spatial variance within the pixels in the ROI (denoted by σ_{ROI}^2) and in the background (denoted by σ_B^2) as follows:

$$\sigma_{ROI}^2 = \frac{1}{R(N_R - 1)}\sum_{k=1}^{R}\sum_{r=1}^{N_R} (x_{r,k} - \theta_{\mathrm{ROI},k})^2. \tag{4.16}$$

$$\sigma_{B}^2 = \frac{1}{R(N_B - 1)}\sum_{k=1}^{R}\sum_{b=1}^{N_B} (x_{b,k} - \theta_{\mathrm{B},k})^2. \tag{4.17}$$

The RMSE over the entire 3D image was computed as below:

$$\mathrm{RMSE} = \frac{1}{R}\sum_{k=1}^{R}\sqrt{\frac{\sum_{i=1}^{N}(x_{i,k} - x_{i,k}^{true})^2}{\sum_{i=1}^{N}(x_{i,k}^{true})^2}} \times 100\%, \tag{4.18}$$

In this and all the other experiments in this work, 100 noise realizations were used to compute the various figures of merit. To study the sensitivity of our method to noise, experiments were conducted with SNR ranging from 5 dB to 40 dB, with step size of 5 dB.

In the second set of experiments, we conducted simulation studies with a digital mouse phantom[220]. Three fluorescence targets were placed in the mouse brain. Two

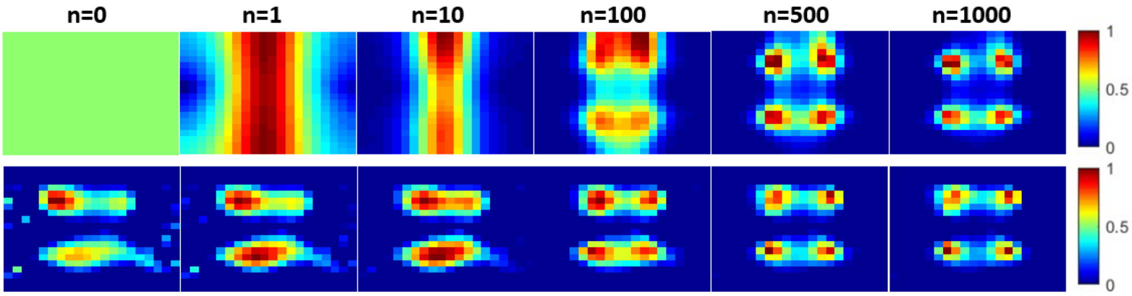


Figure 4-3. Cross sections at $y = 2.5$ cm reconstructed by MLEM with different iteration number n for a cube phantom. SNR=18 dB. The reconstructed images are from MLEM with uniform initial estimate for the top row and MLEM with sparse initial estimate for the bottom row.

of them had a radius of 0.8 mm and the third had a radius of 1.2 mm. The optical properties of the mouse head are listed in Table 4-I. The whole brain was discretized into 2942 voxels. 48 sources and 51 detectors were placed at the surface of the mouse head, as shown in Fig. 4-2(a). The cross section of the phantom at $z = 16$mm is shown in Fig. 4-2(b).

First, 1000 iterations were performed for MLEM with uniform and sparse initialization to study the effect of iteration number on MLEM performance. The SNR was set to 18 dB. The truncation number K was set to 120. Next, quantitative performance of pure sparse reconstruction, sparsity-initialized MLEM and uniformly initialized MLEM methods at different noise levels were evaluated. The SNR value ranged from 5 dB to 40 dB, with step size of 5 dB.

The selection of truncation number plays an important role in the quality of the reconstructed image acquired from sparse reconstruction[38, 69]. For this reason, we also studied the effect of truncation number on reconstruction results of pure sparse reconstruction method and the proposed sparsity-initialized MLEM method. To compare our proposed method and pure sparse reconstruction method, we conducted experiments with different truncation number K. 450 iterations were used for MLEM with sparse initial estimate. For quantitative study, RMSE was computed as a function

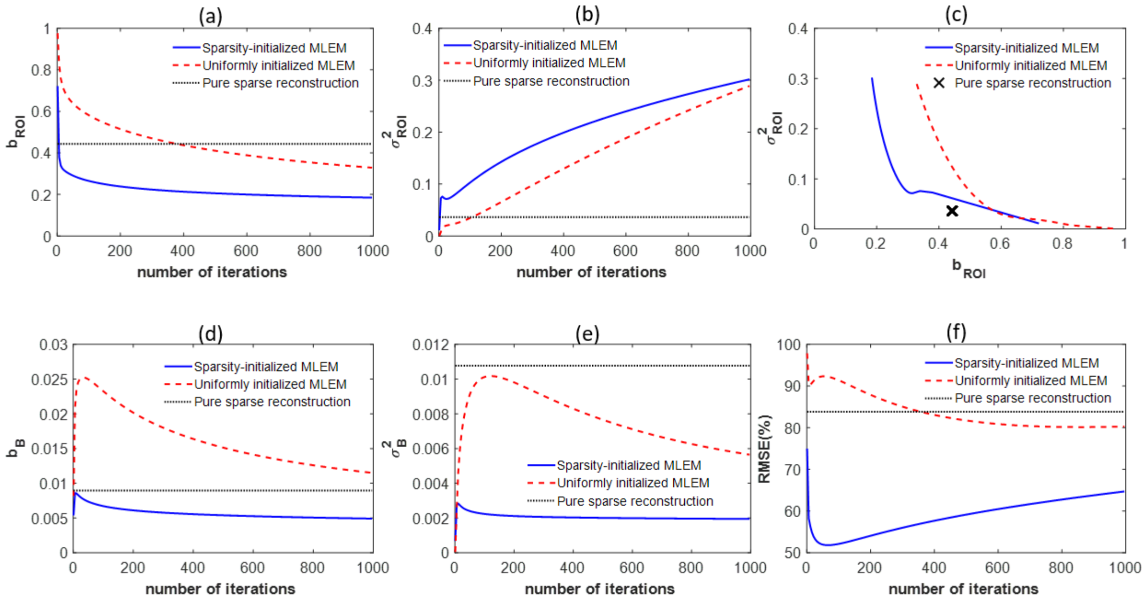


Figure 4-4. Quantitative results of different reconstruction methods as functions of iteration number for cube phantom. (a) Plot of ROI bias vs. number of iterations. (b) Plot of ROI spatial variance vs. number of iterations. (c) Plot of ROI spatial variance vs. ROI bias. (d) Plot of background bias vs. number of iterations. (e) Plot of background variance vs. number of iterations. (f) Plot of RMSE vs. number of iterations.

of the truncation number. The experiments were conducted for two noise levels, namely SNR=40 dB and SNR=20 dB.

4.3 Results

4.3.1 Uniform cube phantom

Fig. 4-3 shows cross sections reconstructed by MLEM with different iteration numbers. For sparsity-initialized MLEM, iteration number $n = 0$ corresponds to the case of pure sparse reconstruction. The fluorescence intensity in all figures were normalized to the range of $[0, 1]$. The computation time required by MLEM with different initializations for 1000 iterations is provided in Table 4-II. It can be observed that sparsity-initialized MLEM is about 8 times faster than uniformly initialized MLEM.

Fig. 4-4 shows the quantitative results as a function of iteration number. We observe from these plots that sparsity-initialized MLEM converges at a lower num-

Table 4-II. Computation time required by MLEM for 1000 iterations.

Method	Time (s)
Sparsity-initialized MLEM	5
Uniformly initialized MLEM	39

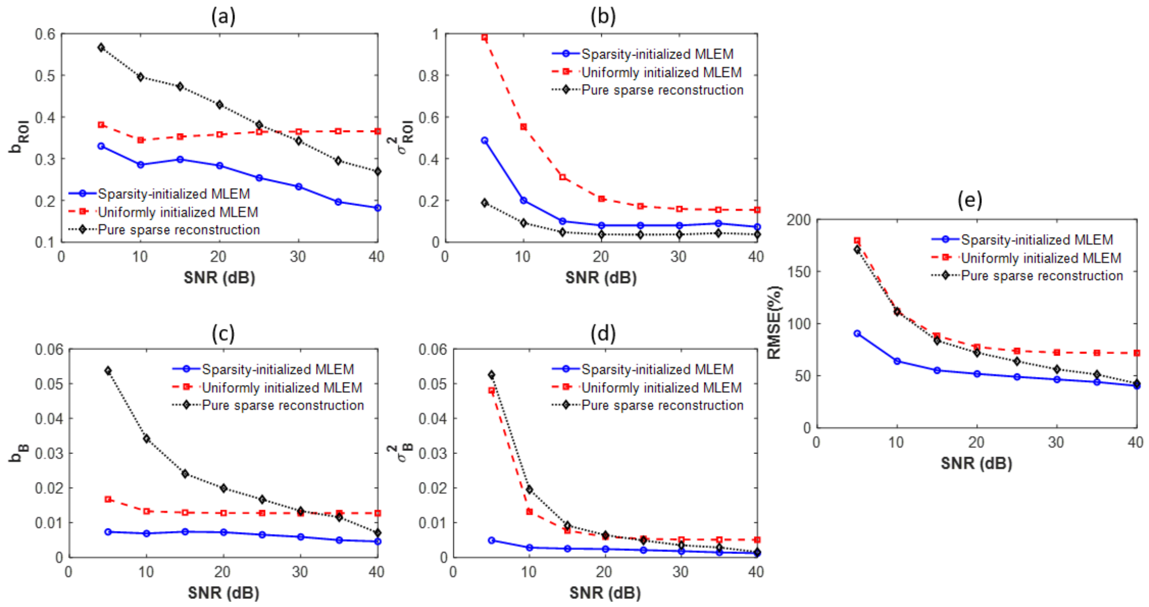


Figure 4-5. Quantitative results of different reconstruction methods as functions of SNR for cube phantom. (a) Plot of ROI bias vs. SNR. (b) Plot of ROI variance vs. SNR. (c) Plot of background bias vs. SNR. (d) Plot of background variance vs. SNR. (e) Plot of RMSE vs. SNR.

ber of iterations. Sparsity-initialized MLEM has lower ROI bias, background bias, background spatial variance, and image RMSE. We also computed the variance of the mean ROI and the mean background uptakes, and found that these were much lower (less than 1%) compared to the bias. Thus, we do not show these results here. From Fig. 4-4(f), we notice that sparsity-initialized MLEM reached its lowest RMSE after only 50 iterations, but for uniformly initialized MLEM, the lowest RMSE was obtained after 800 iterations. Based on this result, we chose 50 iterations for sparsity-initialized MLEM and 800 iterations for uniformly initialized MLEM for the first set of experiments with different SNR values. The plots of quantitative results for the different reconstruction methods at different SNR values are shown in Fig. 4-5.

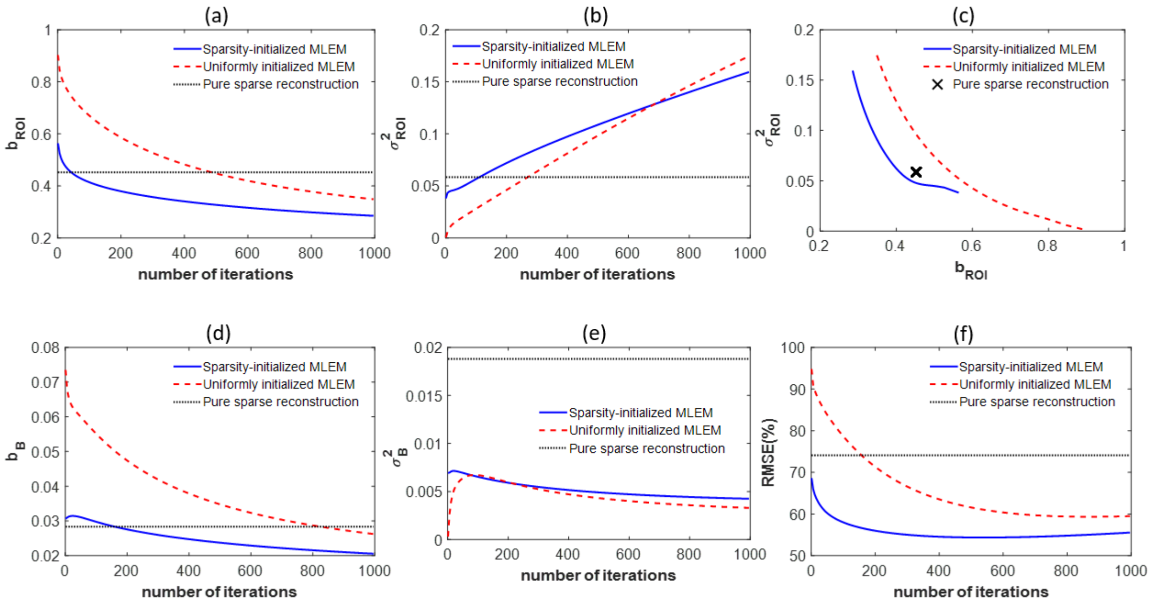


Figure 4-6. Quantitative results of different reconstruction methods as functions of iteration number for digital mouse phantom. (a) Plot of ROI bias vs. number of iterations. (b) Plot of ROI spatial variance vs. number of iterations. (c) Plot of ROI spatial variance vs. ROI bias. (d) Plot of background bias vs. number of iterations. (e) Plot of background variance vs. number of iterations. (f) Plot of RMSE vs. number of iterations.

we again observe that sparsity-initialized MLEM leads to lower ROI bias, background bias, background spatial variance, and image RMSE for all noise levels.

4.3.2 Digital mouse phantom

Fig. 4-6 shows quantitative performance of different reconstruction methods as a function of iteration number. We observe that sparsity-initialized MLEM achieves lower ROI bias, background bias and RMSE. It was observed that for the sparsity-initialized MLEM and uniformly initialized MLEM, 450 and 900 iterations yielded the minimum RMSE. Thus, these values were chosen for the two methods for subsequent experiments. Quantitative performance of different reconstruction methods at different noise levels is shown in Fig. 4-7. The sparsity-initialized MLEM method shows better performance for ROI bias, background bias and RMSE compared to the other two methods.

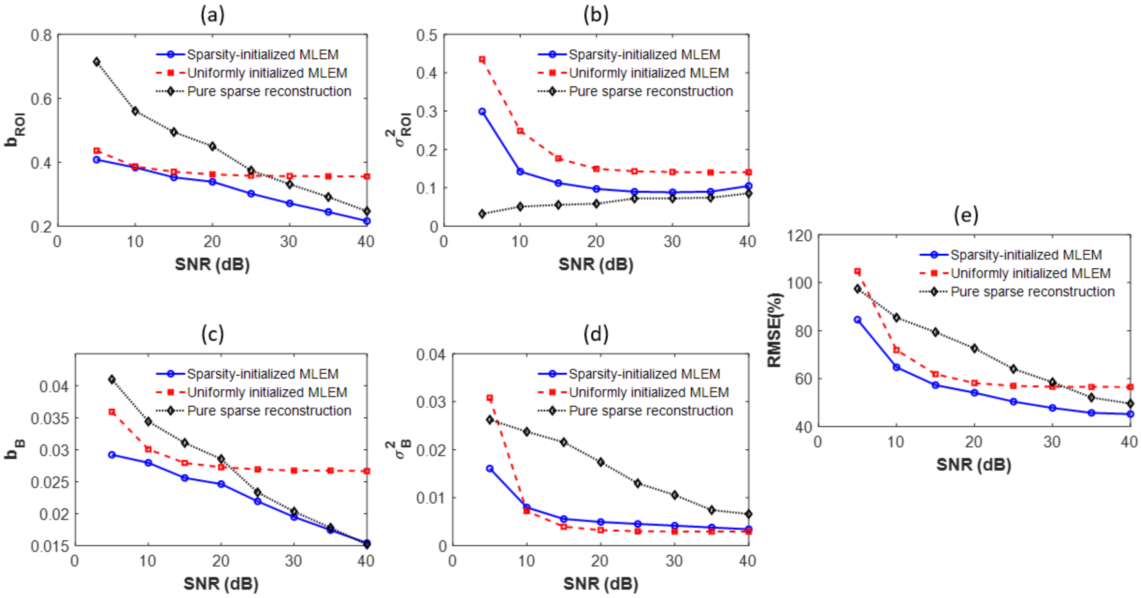


Figure 4-7. Quantitative results of different reconstruction methods as functions of SNR for digital mouse phantom. (a) Plot of ROI bias vs. SNR. (b) Plot of ROI variance vs. SNR. (c) Plot of background bias vs. SNR. (d) Plot of background variance vs. SNR. (e) Plot of RMSE vs. SNR.

Fig. 4-8 shows the cross sections reconstructed by pure sparse reconstruction and MLEM with sparse initial estimate for different truncation number. From Fig. 4-8, we notice that for small truncation number, pure sparse reconstruction generates blurry images. As truncation number increases, the resolution improves, but the background noise also increases due to the amplification of noise during preconditioning. For truncation number larger than 550, the signal is totally overwhelmed by the noise. As a comparison, the proposed method is able to largely reduce the background noise as truncation number increases. The RMSE as a function of truncation number is plotted in Fig. 4-9. The sparsity-initialized MLEM leads to lower RMSE for both noise levels.

4.4 Discussion

In this work, we have proposed an MLEM-based technique to reconstruct the fluorescence distribution from FMT data. In our framework, the initial estimate for the MLEM algorithm is derived from a sparse reconstruction method. Often an uniform initial estimate is used with MLEM-based techniques, but here we observe that a sparsity-initialized technique yields several advantages compared to uniformly initialized MLEM. First, sparsity-initialized MLEM has faster convergence speed. From Table 4-II, Fig. 4-3, Fig. 4-4(a) and Fig. 4-6(a), we observe that sparse initial estimate speeds up the convergence by both shortening the computation time for each iteration and requiring fewer iterations for convergence. In addition, sparsity-initialized MLEM also provides improved quantitative performance in ROI bias, background bias, ROI spatial variance, RMSE, and bias-variance trade-off compared to uniformly initialized MLEM, as shown in Fig. 4-4-4-7. Further, while results in both the cube phantom and the digital mouse phantom experiments indicate that the proposed method leads to higher ROI spatial variance compared to uniformly initialized MLEM for the same number of iterations, Fig. 4-4(c) and Fig. 4-6(c) show that the proposed method still provides better bias-variance trade-off compared to MLEM with uniform initial estimate. Further, sparsity-initialized MLEM often requires fewer iterations, which enables it to provide lower ROI spatial variance compared to uniformly initialized MLEM, as we observe from Fig. 4-5(b) and Fig. 4-7(b).

We also observe that sparsity-initialized MLEM provides advantages over pure sparse reconstruction method. From Fig. 4-8, we notice that sparsity-initialized MLEM is less sensitive to the choice of truncation number. For pure sparse reconstruction method, when the truncation number is small, the reconstructed image is blurry. As truncation number increases, image resolution is improved, but the noise in the background region is also increased due to the noise amplification during preconditioning.

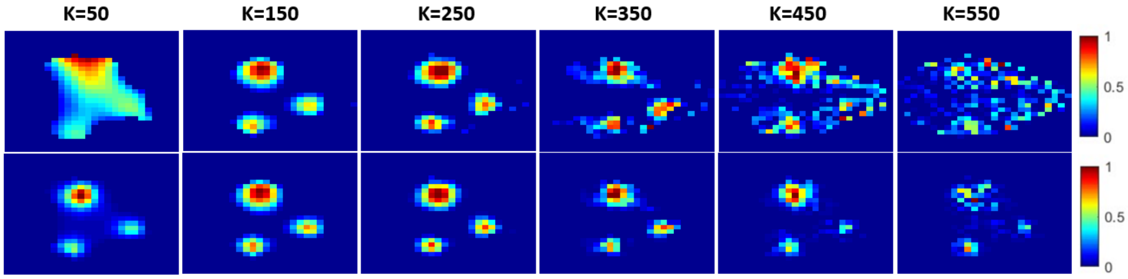
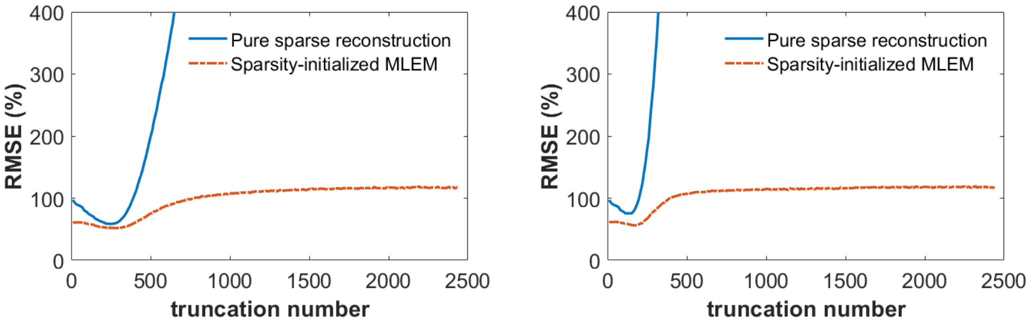


Figure 4-8. Cross sections of fluorescence target reconstructed with pure sparse reconstruction method for the top row and the proposed method for the bottom row with different truncation number K for digital mouse phantom for SNR=40dB.

On the other hand, for small truncation number, sparsity-initialized MLEM is able to improve the resolution compared to pure sparse reconstruction method. For large truncation number, sparsity-initialized MLEM reduces noise in the background. These properties make MLEM with sparse initial estimate more robust to the choice of truncation number compared to pure sparse reconstruction method. The plots of RMSE vs. truncation number in Fig. 4-9 also demonstrate this point. Sparsity-initialized MLEM also improves quantitative performance of reconstructed images compared to pure sparse reconstruction method. Fig. 4-5 and Fig. 4-7 indicate this for different SNR values. Apart from improved background bias and spatial variance due to the reduction of background noise, sparsity-initialized MLEM also reduces the ROI bias compared to pure sparse reconstruction method, especially at low SNR value. At low SNR value, small truncation number is preferred to avoid noise amplification, which results in only a small number of measurements used for reconstruction. Small truncation number not only generates blurry images, as we discussed previously, but also causes severe bias in the reconstruction results. For example, for SNR= 5 dB, we observe that pure sparse reconstruction generated 71% ROI bias in the cube phantom experiments and 57% ROI bias in the digital mouse phantom experiments. As a comparison, MLEM uses the original system matrix and detector measurements for reconstruction, enabling it to compensate for the bias in the image, which reduced



(a)

(b)

Figure 4-9. Plot of RMSE vs. truncation number for pure sparse reconstruction method and the proposed reconstruction method for different noise levels. (a) Plot of RMSE vs. truncation number for SNR=40 dB. (b) Plot of RMSE vs. truncation number for SNR=20 dB.

the ROI bias to 40% in the cube phantom experiments and 33% in the digital mouse phantom experiments.

We have observed, for example in Fig. 4-3, that sparsity-initialized MLEM is able to suppress the noise in the background region that is present in the sparse initial estimate. To explain this observation, here we provide a theoretical justification. For a set of detector measurements denoted by $\boldsymbol{\Phi}$, consider two reconstructed images $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, where $\boldsymbol{x}_1$ is image with noise in the background region (referred to as background noise), and $\boldsymbol{x}_2$ is image that does not contain this background noise, as shown in Fig. 4-10(a) and (b), respectively. We denote the background noise as $\epsilon = \boldsymbol{x}_1 - \boldsymbol{x}_2$, where $\epsilon_n \geq 0$ for all n. Before we proceed further, we introduce the concept of KL distance. This distance measures how two probability distributions diverge from another. It is known that MLEM attempts to find an estimate that minimizes the Kullback-Leibler (KL) distance between the measured data $\boldsymbol{\Phi}$ and the data predicted by an estimate $\mathbf{G}\boldsymbol{x}$. Thus, our objective is to assess whether the KL distance of $\boldsymbol{x}_2$ is less than $\boldsymbol{x}_1$, which would explain why MLEM would yield a solution $\boldsymbol{x}_2$ in comparison to $\boldsymbol{x}_1$.

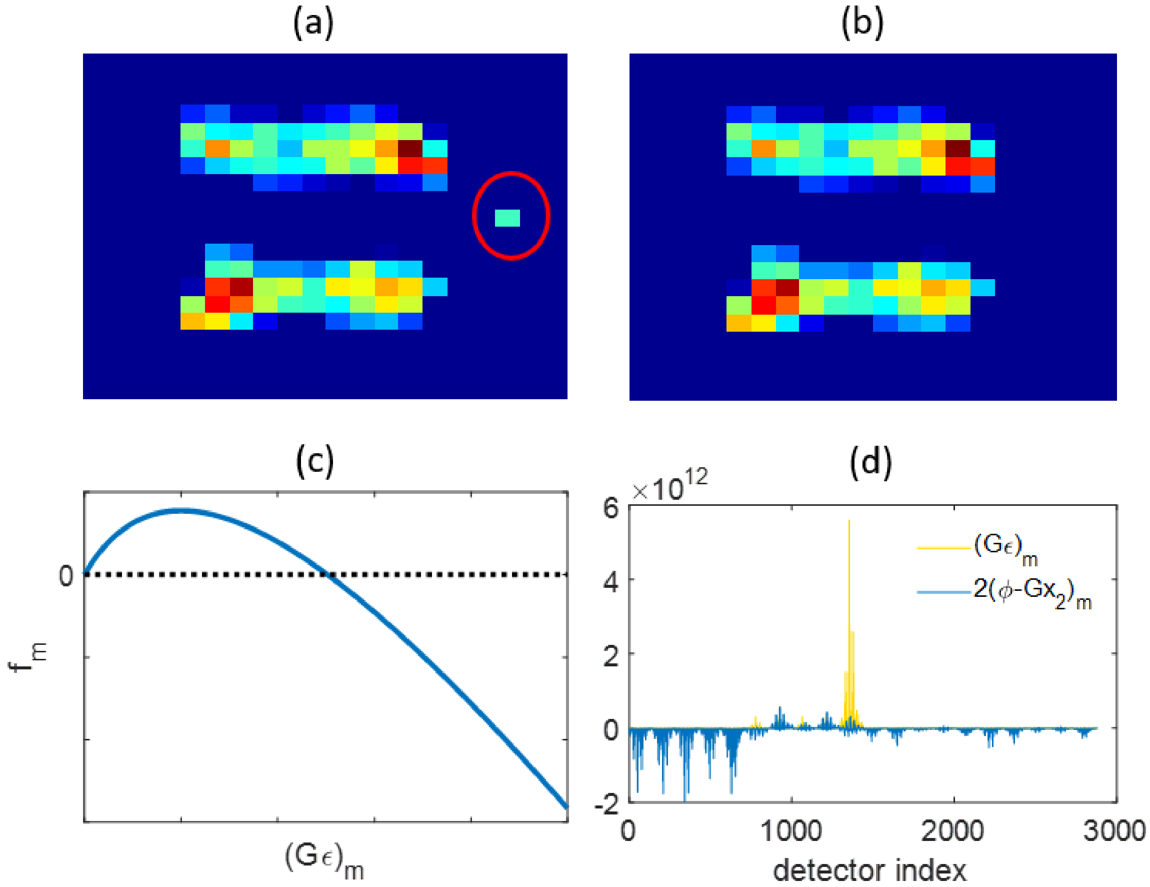


Figure 4-10. (a) Image with background noise. The noise spot in the background is marked with red circle. (b) Image without background noise. (c) Plot of f_m. (d)Plot of $(\mathbf{G}\epsilon)_m$ and $2(\phi - \mathbf{G}\boldsymbol{x}_2)_m$.

For $\boldsymbol{x}_1$, the KL distance is:

$$D_{KL,1}(\mathbf{\Phi}, \mathbf{G}\boldsymbol{x}_1) = \sum_m \left\{ (\mathbf{G}\boldsymbol{x}_1)_m - \phi_m + \phi_m \ln \frac{\phi_m}{(\mathbf{G}\boldsymbol{x}_1)_m} \right\}. \tag{4.19}$$

Similarly, for $\boldsymbol{x}_2$, the KL distance is:

$$D_{KL,2}(\mathbf{\Phi}, \mathbf{G}\boldsymbol{x}_2) = \sum_m \left\{ (\mathbf{G}\boldsymbol{x}_2)_m - \phi_m + \phi_m \ln \frac{\phi_m}{(\mathbf{G}\boldsymbol{x}_2)_m} \right\}. \tag{4.20}$$

We then have the difference:

$$\begin{aligned} \Delta D_{KL} &= D_{KL,2}(\mathbf{\Phi}, \mathbf{G}\boldsymbol{x}_2) - D_{KL,1}(\mathbf{\Phi}, \mathbf{G}\boldsymbol{x}_1) \\ &= \sum_m \left\{ -(\mathbf{G}\epsilon)_m + \phi_m \ln \left[1 + \frac{(\mathbf{G}\epsilon)_m}{(\mathbf{G}\boldsymbol{x}_2)_m} \right] \right\}. \end{aligned} \tag{4.21}$$

We denote $f_m((\mathbf{G}\epsilon)_m) = -(\mathbf{G}\epsilon)_m + \phi_m \ln \left[1 + \frac{(\mathbf{G}\epsilon)_m}{(\mathbf{G}\boldsymbol{x}_2)_m} \right]$. If $\phi_m \leq (\mathbf{G}\boldsymbol{x}_2)_m$, $f_m \leq 0$ since f_m is monotonically decreasing for $(\mathbf{G}\epsilon)_m \leq 0$ and $f_m(0) = 0$. If $\phi_m > (\mathbf{G}\boldsymbol{x}_2)_m$, the plot of f_m is shown in Fig. 4-10(c). To estimate the zeros of f_m, we use second order Taylor expansion to approximate f_m, which gives:

$$f_m \approx -(\mathbf{G}\epsilon)_m + \phi_m \left[\frac{(\mathbf{G}\epsilon)_m}{(\mathbf{G}\boldsymbol{x}_2)_m} - \frac{1}{2} \left(\frac{(\mathbf{G}\epsilon)_m}{(\mathbf{G}\boldsymbol{x}_2)_m} \right)^2 \right]. \tag{4.22}$$

Let $f_m = 0$, we have

$$(\mathbf{G}\epsilon)_{m,1} = 0, \tag{4.23}$$

and

$$\begin{aligned} (\mathbf{G}\epsilon)_{m,2} &= \frac{2(\mathbf{G}\boldsymbol{x}_2)_m[\phi_m - (\mathbf{G}\boldsymbol{x}_2)_m]}{\phi_m} \\ &\leq 2[\phi_m - (\mathbf{G}\boldsymbol{x}_2)_m], \end{aligned} \tag{4.24}$$

where the inequality comes from the fact that $\phi_m > (\mathbf{G}\boldsymbol{x}_2)_m$. Also, note that the function f_m has its maxima at $(\mathbf{G}\epsilon)_m = \phi_m - (\mathbf{G}\boldsymbol{x}_2)_m$. Thus, if the detector response to the noise spot has similar pattern as $\phi_m - (\mathbf{G}\boldsymbol{x}_2)_m$, f_m will be close to its maximum for most detector index m. This provides a higher chance that $\Delta D_{KL} > 0$, which means MLEM is more likely to update towards noisy image. This is the case for the noise close to ROI. On the other hand, for noise spot in background region, the detector response to noise spot will have a very different pattern compared to $\phi_m - (\mathbf{G}\boldsymbol{x}_2)_m$, as shown in Fig. 4-10(d). For detector index m where $\phi_m - (\mathbf{G}\boldsymbol{x}_2)_m > 0$, $(\mathbf{G}\epsilon)_m$ is either close to 0 or too large. This results in makes f_m either close to zero or have a negative value. In this case, it has higher chance that $\Delta D_{KL} < 0$, meaning MLEM tends to update towards results without the noise spot.

The noise model in FMT is often assumed to be Gaussian [38, 72, 83, 196, 197, 217]. In very few cases is the Poisson noise model applied [201]. Gaussian noise model is a good approximation when SNR is high, i.e. sufficient number of photons are detected. However, in some applications, the SNR value might be low, such as in brain imaging [26], dynamic FMT [221] and early-photon FMT [43]. Our results demonstrate that incorporating the Poisson noise model is especially valuable in these scenarios. More specifically, the pure sparse reconstruction method was formulated based on Gaussian noise model, while the proposed method incorporated both the sparsity information and Poisson noise model. We observe that the performance of the proposed method improves in comparison to the pure sparse reconstruction method as the SNR value decreases, and the proposed method is substantially more reliable at low SNR values. This shows the importance of accurately modeling Poisson noise for applications of FMT when insufficient number of photons are detected. We also point out that there might be some instances where the data is not a pure Poisson distribution. This could occur, for example in case of CMOS detectors where the noise is a combination of Poisson and Gaussian distribution and is effected by electronic gain [222]. However, even in those cases, MLEM provides a convenient way to account for non-negativity constraints and enforce photon conservation. These factors may help to improve the quality of reconstruction in comparison to just using a sparsity-based method. Thus, exploring the performance of this method with systems that have non-Poisson noise distribution would be another important future study.

In this work, we only considered the case where the background uptake of fluorescence distribution is zero. While this is a common assumption in FMT studies [38, 69, 72, 83, 196, 197, 217], it is possible that the background uptake is non-zero. Exploring the performance of the proposed method for this task would be an important future direction. The proposed method has been validated with extensive simulation experiments. Evaluating the performance of the method with physical phantom and *in vivo* animal experiments is another important direction of research. Finally, we used the MC-based method to model photon propagation to obtain the Greens function in this work. However, there have been several analytical methods proposed for modeling light transport[12, 54, 55, 59, 184, 187]. These methods can also be used to obtain an expression for the Green's function. Analytical methods offer the advantage that they might be less sensitive to photon noise. Thus, implementing this reconstruction method using the analytical approaches is another important research direction.

4.5 Conclusion

We have presented a reconstruction framework for FMT involving sparsity-initialized MLEM. Simulation experiments on cubic digital mouse phantoms demonstrate that the proposed method yields improved qualitative and quantitative performance compared to uniformly initialized MLEM as well as sparsity-initialized MLEM techniques. Further, compared to uniformly initialized MLEM, the proposed method is faster to execute, overcoming another barrier to application of MLEM technique for optical tomography. Moreover, compared to pure sparse reconstruction, the proposed method is more robust to noise amplification. We have also provided theoretical justification for the ability of the proposed method to reduce noise in the background region. Overall, this work provides strong evidence that the proposed sparsity initialized MLEM-based reconstruction framework is feasible and advantageous for reconstruction in FMT imaging systems.

Chapter 5

Partial volume correction for brain PET imaging

In this chapter, we first introduce commonly used PVC methods. We then present a post-reconstruction PVC method based on deconvolution with parallel level set (PLS) regularization. We frame the problem as an iterative deconvolution task with PLS regularization that incorporates anatomical information without requiring MR segmentation or assuming uniformity of PET distributions within regions. An efficient algorithm for the non-smooth optimization of the objective function is developed so that the proposed method can be feasibly applied for 3D images and produces sharper images compared to conventional PLS method with smooth optimization. The proposed method was evaluated with both simulated data and *in vivo* human data collected from Baltimore Longitudinal Study of Aging (BLSA). Results indicate our method is able to reduce image noise while preserving structure details, as well as increasing the number of statistically significant voxels. Further, our statistical analysis shows that our method has the potential to better differentiate amyloid positive and amyloid negative scans. Overall, our results demonstrate promise to provide superior performance in clinical imaging scenarios.

5.1 Introduction

Due to the limited spatial resolution of PET systems, the reconstructed PET images suffer from partial volume effect (PVE), as we have discussed in 2.2.5. For brain PET imaging, structures such as cortical gray matter may have a thickness of only 4 mm and can be severely influenced by PVE. The situation becomes even worse when studying neurodegenerative disease such as Alzheimer's disease and Parkinson disease, where atrophy occurs in different structures [223, 224].

In order to correct such effect, different partial volume correction (PVC) algorithms have been developed. These methods can be roughly categorized into reconstruction-based methods and post-reconstruction-based methods [122, 225]. For reconstruction-based methods, resolution degradation factors are modeled in system matrix. There are different ways to model these factors in system matrix [151, 226–232]. Here we give an example of resolution modeling based on the following matrix factorization [151, 230, 232]:

$$\boldsymbol{P} = \boldsymbol{P}_{det,sens}\boldsymbol{P}_{det,blur}\boldsymbol{P}_{attn}\boldsymbol{P}_{geom}\boldsymbol{P}_{blur,im}. \tag{5.1}$$

In this factorization, $\boldsymbol{P}_{geom}$ represents the forward projection matrix. $\boldsymbol{P}_{attn}$ and $\boldsymbol{P}_{det,sens}$ contains the attenuation and normalization factors, respectively. Resolution degradation factors are modeled through $\boldsymbol{P}_{det,blur}$ and $\boldsymbol{P}_{blur,im}$. Specifically, $\boldsymbol{P}_{det,blur}$ accounts for sinogram blurring factors including photon-pair noncollinearity, inter-crystal scatter and penetration. $\boldsymbol{P}_{blur,im}$ models image space blurring factors such as positron range. Direct reconstruction with point spread function (PSF) modeling may cause Gibbs artifacts near edges. To alleviate such artifact, Bayesian MAP reconstruction can be applied to incorporate prior information [233].

On the other hand, post-reconstruction-based methods usually model PVE by assuming the observed image is a result of the true image convolved with PSF. Inversion of this process is achieved through either voxel-based methods or region-of-interest-based methods. Compared with reconstruction-based PVC methods, post-reconstruction methods do not require access to image raw data or scanner reconstruction algorithm, and as such are more applicable to large scale (especially multi-center) studies.

We now briefly introduce some commonly used post-reconstruction-based methods. Let Ω be a cubic domain in $\mathbb{R}^3$. $g, u : \Omega \to \mathbb{R}$ denote the observed (PVE contaminated) and true PET images, respectively. h represent system PSF. Most post-reconstruction PVC methods use high-resolution anatomical information from MRI. Further, many PVC methods require segmenting MRI into different tissues. Assume we segment Ω into $\Omega_1 \cdots \Omega_M$ such that $\Omega = \Omega_1 \cup \cdots \cup \Omega_M$. We define a mask function for each segmentation region as

$$m_i(\boldsymbol{r}) = \begin{cases} 1, \text{if } \boldsymbol{r} \in \Omega_i. \\ 0, \text{otherwise} \end{cases} \tag{5.2}$$

Geometric transfer matrix method

Rousset et al. [234] developed a so-called geometric transfer matrix (GTM) method for PVC. In this method, the spill-over effect between region i and region j is addressed by computing a weighting factor as

$$w_{ij} = \frac{1}{n_j} \int_{\Omega_j} \mathrm{RSF}_i(\boldsymbol{r}) d\boldsymbol{r}, \tag{5.3}$$

where n_i is the number of pixels in region j, $\mathrm{RSF}_i(\boldsymbol{r})$ is regional spread function and is computed as $\mathrm{RSF}_i(\boldsymbol{r}) = (h * m_i)(\boldsymbol{r})$. By further assuimg the observed image g and the true image u are piecewise-constant images such that $g(\boldsymbol{r}) = \sum_i t_i m_i(\boldsymbol{r})$ and $u(\boldsymbol{r}) = \sum_i T_i m_i(\boldsymbol{r})$, where t_i and T_i are regional mean activities in g and u, respectively. Then t_i and T_i are linked with

$$\boldsymbol{t} = \boldsymbol{W}\boldsymbol{T}, \tag{5.4}$$

where $\boldsymbol{t} = [t_1, \cdots, t_M]^T$, $\boldsymbol{T} = [T_1, \cdots, T_M]^T$ and w_{ij} computed in 5.3 are elements in $\boldsymbol{W}$. Corrected regional mean value $\boldsymbol{T}$ can be obtained with

$$\boldsymbol{T} = \boldsymbol{W}^{-1}\boldsymbol{t}. \tag{5.5}$$

Labbé et al. [235] developed a similar method. In Labbé's method, the observed image $\boldsymbol{g}$ is not assumed to be piecewise-constant, thus the weighting matrix $\boldsymbol{W}$ addresses the spill-over effect between ROIs in the true image and voxels in the observed image. Sattarivand et al. [236] developed a so-called symmetric GTM method which is mathematically equivalent to Labbé's method but reduces the computational burden. Du et al. [237] proposed a perturbation-based GTM method that took nonlinearity of the reconstruction process into account. The GTM-type methods recover corrected ROI mean activities of the true image, but they do not produce voxel-based images.

Müller-Gärtner's method

In Müller-Gärtner's method [238], the true brain image is treated as the summation of three compartment: gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF). This can be represented as

$$g(\boldsymbol{r}) = (h * u)(\boldsymbol{r}) = h(\boldsymbol{r}) * \left[(u_{GM} m_{GM})(\boldsymbol{r}) + (u_{CSF} m_{CSF})(\boldsymbol{r}) + (u_{WM} m_{WM})(\boldsymbol{r})\right], \tag{5.6}$$

where u_{GM}, u_{WM} and u_{CSF} are images of GM, WM, and CSF, respectively. By assuming WM and CSF have constant activities such that $u_{WM}(\boldsymbol{r}) = T_{WM}$ and $u_{CSF}(\boldsymbol{r}) = T_{CSF}$, the corrected GM image is obtained with

$$u_{GM}(\boldsymbol{r}) = \frac{g(\boldsymbol{r}) - h(\boldsymbol{r}) * (T_{WM} m_{WM})(\boldsymbol{r}) - h(\boldsymbol{r}) * (T_{CSF} m_{CSF})(\boldsymbol{r})}{h(\boldsymbol{r}) * m_{GM}(\boldsymbol{r})}. \tag{5.7}$$

In practice, the mean activity T_{WM} and T_{CSF} can be estimated from the observed image g by computing the mean concentrations in smaller ROIs in WM and CSF.

The Müller-Gärtner's method was further extended to four compartments [239] and later was generalized to multiple regions [240, 241]

Yang's method

Yang et al. [242] presented a method where PVC is applied to the whole image. In Yang's method, a simulated piece-wise constant PET image $u_{simulated}$ is created from the segmented MR image. Each region in $u_{simulated}$ is filled with its true relative mean value. Then image after correction is computed as

$$u_{corrected}(\boldsymbol{r}) = g(\boldsymbol{r}) \times \frac{u_{simulated}(\boldsymbol{r})}{(h * u_{simulated})(\boldsymbol{r})} \tag{5.8}$$

Erlandsson et al. [156] improved this method by updating the mean activity estimation in $u_{simulated}$ iteratively, which gives the so-called iterative Yang method. Thomas et al. [243] proposed region-based voxel-wise (RBV) method based on Yang's method that uses GTM method to estimate $u_{simulated}$.

Iterative deconvolution method

There are two classical deconvolution methods: reblurred Van-Cittert (rVC) method and Richardson-Lucy (RL) method, that are based on Gaussian and Poisson noise models, respectively [244]. In rVC method, we aim to solve the following least-square minimization problem:

$$\min_{u} \|g(\boldsymbol{r}) - h(\boldsymbol{r}) * u(\boldsymbol{r})\|_2^2, \tag{5.9}$$

where $\|\cdot\|_2$ denotes L_2 norm. The steepest descent update for the above equation is

$$u^{(k+1)}(\boldsymbol{r}) = u^{(k)}(\boldsymbol{r}) + \alpha h(\boldsymbol{r}) * [g(\boldsymbol{r}) - h(\boldsymbol{r}) * u^{(k)}(\boldsymbol{r})], \tag{5.10}$$

where the step length α is typically set as $0 < \alpha < 2$. Update in 5.10 is referred to as rVC method. RL deconvolution has a similar update equation as MLEM but the system matrix is replaced with a convolution operator with system PSF.

For the above-mentioned MR-guided PVC methods, we notice that commonly uniform PET activity distribution is assumed within a given anatomical region, as determined by the segmented MRI. However, this assumption is often violated in reality, which may influence the performance of these methods. On the other hand, for iterative deconvolution methods, a known issue is that they lead to noise amplification as the number of iterations increases [245]. To overcome this issue, different noise suppression techniques either through filtering or regularization have been incorporated with the deconvolution framework. For example, in [246], a regularized deconvolution method with adaptive regularization parameter based on local topology was proposed. in [247], wavelet denoising is applied on data residual during deconvolution process to control noise. In [248], non-local-mean-based regularization is incorporated into deconvolution to provide anatomical guidance as well as controlling noise.

In the present work, we develop a post-reconstruction PVC method that relies on 'subtle' MR guidance, meaning that the method does not assume uniform PET activity distribution and also does not force strong boundaries in the PET images based solely on boundary information provided by anatomical guidance. Deconvolution-based methods provide an excellent framework to minimize or avoid such assumptions on PET activity distributions. In order to incorporate anatomical information to suppress noise while further improving image resolution, here we choose the parallel level set (PLS) function as the regularization term. PLS was proposed recently and has been used for PET reconstruction and PET-MRI joint reconstruction [249, 250]. PLS allows incorporating anatomical information without segmentation and reduces to total variation when local MRI map is uniform, allowing it to preserve edges in PET when corresponding structure is missing in MRI. On the other hand, unlike functions such as total variation that does not use any anatomical guidance, PLS allows recovery of structures in PET wherever boundary matches with that from anatomical guidance. This provides images with higher resolution compared to those recovered without using anatomical guidance. Overall, we propose a post-reconstruction PVC method based on regularized deconvolution with PLS regularization. The proposed method allows performing anatomical-guided PVC without assumption of uniformity on PET activity distribution and segmentation of MR image. Moreover, conventional PLS requires adding a smoothing factor so that gradient-based optimization methods are applicable. By contrast, in this work, we develop a non-smooth optimization strategy based on split Bregman method to minimize the objective function of the regularized deconvolution problem, which enables applying the method on 3D image in a computationally efficient manner while avoiding blur caused by smoothing. The proposed method was evaluated and compared against other post-reconstruction PVC methods using both simulation data and clinical data.

5.2 Methods

5.2.1 Regularized Deconvolution Model

We have defined g, u as the observed and true PET images, and h as system PSF. The regularized deconvolution model is given as

$$u^{*} = \arg\min_{u} \frac{\mu}{2}\|g - h * u\|_2^2 + R(u), \tag{5.11}$$

where $R(u)$ represents regularization function, μ is a hyper-parameter that balances deconvolution and regularization.

5.2.2 Parallel Level Sets Regularization

In this work, we utilize PLS as the regularization function in (5.11). PLS function is defined as [250]:

$$R_{PLS}(u) = \int_{\Omega} \left(\beta^2 + \|\nabla u(\boldsymbol{r})\|_2^2 - \langle \nabla u(\boldsymbol{r}), \xi(\boldsymbol{r})\rangle^2\right)^{\frac{1}{2}} d\boldsymbol{r}, \tag{5.12}$$

where $\|\nabla u(\boldsymbol{r})\|_2^2 = \left[(\nabla_x u(\boldsymbol{r}))^2 + (\nabla_y u(\boldsymbol{r}))^2 + (\nabla_z u(\boldsymbol{r}))^2\right]^{1/2}$ is the L_2 norm for image gradient of PET, $\langle \cdot, \cdot \rangle$ represents the inner product, $\xi(\boldsymbol{r}) = \nabla v(\boldsymbol{r})/(\eta^2 + \|\nabla v(\boldsymbol{r})\|_2^2)^{\frac{1}{2}}$ denotes the normalized MR gradient with $v : \Omega \to \mathbb{R}$ representing MR image, and β and η are two smoothing factors.

5.2.3 Optimization Algorithm

When $\beta \neq 0$ in (5.12), the PLS regularization function is differentiable and thus, the minimization problem in (5.11) can be solved using gradient-based methods. However, the smoothing factor β may lead to undesirable blur in the output images. In this section we applied split Bregman method [251] to provide a non-smooth optimization strategy for our problem.

It was shown in [250] that

$$\|\nabla u(\boldsymbol{r})\|_2^2 - \langle \nabla u(\boldsymbol{r}), \xi(\boldsymbol{r})\rangle^2 = \|\boldsymbol{B}(\boldsymbol{r})\nabla u(\boldsymbol{r})\|_2^2, \tag{5.13}$$

where $\boldsymbol{B}(\boldsymbol{r}) = (\boldsymbol{I} - c\xi\xi^T)(\boldsymbol{r})$, $c = \left(1 + [1 - \|\xi(\boldsymbol{r})\|_2^2]^{1/2}\right)^{-1}$ is a normalization constant. When $\beta = 0$, the PLS regularization function can be written as:

$$R_{PLS}(u) = \int_\Omega \|\boldsymbol{B}(\boldsymbol{r})\nabla u(\boldsymbol{r})\|_2. \tag{5.14}$$

For discrete image, if we use $\boldsymbol{g}$, $\boldsymbol{u}$ to denote vectorized observed and true PET images, and replace regularization function in (5.11) with (5.14), we obtain:

$$\min_{\boldsymbol{u}} \frac{\mu}{2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{u}\|_2^2 + \sum_{i=1}^{N} \|\boldsymbol{B}_i(\nabla \boldsymbol{u})_i\|_2, \tag{5.15}$$

where $\boldsymbol{H}$ is blurring matrix with each column $\boldsymbol{h}_i$ representing PSF centered at voxel i, N is the total number of voxels in PET image, $\nabla \boldsymbol{u}$ represents 3D image gradient and can be obtained with finite difference method. If we introduce an auxiliary variable $\boldsymbol{d} = [d_x, d_y, d_z]^T$, the minimization problem in (5.15) can be written as:

$$\min_{\boldsymbol{u}} \frac{\mu}{2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{u}\|_2^2 + \sum_{i=1}^{N} \|\boldsymbol{d}_i\|_2 \quad \text{such that} \quad \boldsymbol{d}_i = \boldsymbol{B}_i(\nabla \boldsymbol{u})_i. \tag{5.16}$$

Let us write (5.16) in an unconstrained form, and we have:

$$\min_{\boldsymbol{u},\boldsymbol{d}} \frac{\mu}{2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{u}\|_2^2 + \sum_{i=1}^{N} \|\boldsymbol{d}_i\|_2 + \frac{\lambda}{2}\sum_{i=1}^{N} \|\boldsymbol{d}_i - \boldsymbol{B}_i(\nabla \boldsymbol{u})_i\|_2^2, \tag{5.17}$$

where λ is another regularization parameter. Conventionally, solution of (5.17) only approximate that of (5.16) when $\lambda \rightarrow \infty$. Bregman iteration guarantees the solution of (5.17) converges to that of (5.16) as the number of iteration increases. To apply Bregman iteration, we update the following two steps iteratively:

$$\min_{\boldsymbol{u},\boldsymbol{d}} \frac{\mu}{2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{u}\|_2^2 + \sum_{i=1}^{N} \|\boldsymbol{d}_i\|_2 + \frac{\lambda}{2}\sum_{i=1}^{N} \|\boldsymbol{d}_i - \boldsymbol{B}_i(\nabla \boldsymbol{u})_i - \boldsymbol{b}_i^{(k)}\|_2^2, \tag{5.18}$$

$$\boldsymbol{b}^{(k+1)} = \boldsymbol{b}^{(k)} + \boldsymbol{B}\nabla \boldsymbol{u}^{(k+1)} - \boldsymbol{d}^{(k+1)}, \tag{5.19}$$

where the superscripts $(k+1),(k)$ here denote iteration number. Solving (5.18) by minimizing $\boldsymbol{u}$ and $\boldsymbol{d}$ separately arrives at a three-step split Bregman iteration process:

$$\boldsymbol{u}^{(k+1)} = \arg\min_{\boldsymbol{u}} \frac{\mu}{2}\|\boldsymbol{g}-\boldsymbol{H}\boldsymbol{u}\|_2^2 + \frac{\lambda}{2}\sum_{i=1}^{N}\|\boldsymbol{d}^{(k)}-\boldsymbol{B}\nabla\boldsymbol{u}-\boldsymbol{b}^{(k)}\|_2^2, \tag{5.20}$$

$$\boldsymbol{d}^{(k+1)} = \arg\min_{\boldsymbol{d}} \sum_{i=1}^{N}\left(\frac{\lambda}{2}\|\boldsymbol{d}_i-\boldsymbol{B}_i(\nabla\boldsymbol{u}^{(k)})_i-\boldsymbol{b}_i^{(k)}\|_2^2+\|\boldsymbol{d}_i\|_2\right), \tag{5.21}$$

$$\boldsymbol{b}^{(k+1)} = \boldsymbol{b}^{(k)} + \boldsymbol{B}\nabla\boldsymbol{u}^{(k+1)} - \boldsymbol{d}^{(k+1)}. \tag{5.22}$$

The $\boldsymbol{u}$ subproblem in (5.20) is differentiable and can be solved with gradient-based methods, such as steepest descent method and conjugate gradient method. Usually these algorithms are applied iteratively until full convergence is achieved. For split Bregman method, solving the subproblem until full convergence is not necessary since the added precision brought about by full convergence will be wasted in the update of the Bregman parameter $\boldsymbol{b}$ [251]. In addition, updating the $\boldsymbol{u}$ subproblem to full convergence could be time-consuming, especially for applications to 3D images. We seek an approximate solution to the $\boldsymbol{u}$ subproblem with as fewer iterations as possible. For this purpose, we applied a one-step gradient descent method that uses Barzilai-Borwein step with backtracking [252, 253]. The algorithm enables approximating the solution with sufficient accuracy with only a single iteration. The $\boldsymbol{d}$ subproblem in (5.21) is nondifferentiable but can be efficiently solved with generalized soft-thresholding [254].

5.2.4 Parameter selection

There are three hyperparameters λ, η and μ in this algorithm: (i) Parameters μ balances deconvolution and regularization strengths. (ii) η controls smoothing added on MR image. These two parameters need to be selected based on noise levels in PET and MR images; (iii) The other parameter λ influences the convergence of the algorithm. We adopted residual balancing method [255] to update this parameter adaptively.

In residual balancing, we first define primal residual $\boldsymbol{m}$ and dual residual $\boldsymbol{n}$ as:

$$\boldsymbol{m}^{(k+1)} = \boldsymbol{B}\nabla \boldsymbol{u}^{(k+1)} - \boldsymbol{d}^{(k+1)}, \tag{5.23}$$

$$\boldsymbol{n}^{(k+1)} = \lambda^{(k)}(\boldsymbol{B}\nabla)^T(\boldsymbol{d}^{(k+1)} - \boldsymbol{d}^{(k)}), \tag{5.24}$$

where $(\boldsymbol{B}\nabla)^T$ denotes adjoint operator of $\boldsymbol{B}\nabla$. Then λ is updated as

$$\lambda^{(k+1)} = \begin{cases} \tau\lambda^{(k)} & \text{if} \|\boldsymbol{m}_{norm}^{(k)}\|_2 > \epsilon\|\boldsymbol{n}_{norm}^{(k)}\|_2, \\ \tau^{-1}\lambda^{(k)} & \text{if} \|\boldsymbol{n}_{norm}^{(k)}\|_2 > \epsilon\|\boldsymbol{m}_{norm}^{(k)}\|_2, \\ \lambda^{(k)} & \text{otherwise,} \end{cases} \tag{5.25}$$

where τ and ϵ are two constants and are typically chosen as $\tau = 2$ and $\epsilon = 10$, $\boldsymbol{m}_{norm}$ and $\boldsymbol{n}_{norm}$ are normalized primal residual and dual residual terms, respectively, which are defined as:

$$\boldsymbol{m}_{norm} = \frac{\boldsymbol{m}}{\max\{\|\boldsymbol{B}\nabla \boldsymbol{u}^{(k+1)}\|_2, \|\boldsymbol{d}^{(k+1)}\|_2\}}, \qquad \boldsymbol{n}_{norm} = \frac{\boldsymbol{n}}{\|(\boldsymbol{B}\nabla)^T\boldsymbol{b}^{(k+1)}\|_2}. \tag{5.26}$$

5.2.5 Summary of the Algorithm

Our framework can be summarized by the following algorithms 1 and 2:

Algorithm 1 (One-step gradient descent algorithm)

```
Input: $\boldsymbol{u}^{(k)}, \boldsymbol{u}^{(k-1)}, C^{(k)}$
Initialize:
        $\rho = 0.4$, $\sigma = 10^{-4}$
        $\boldsymbol{p} = \boldsymbol{u}^{(k)} - \boldsymbol{u}^{(k-1)}$, $\boldsymbol{y} = \partial F^{(k)}(\boldsymbol{u}^{(k)}) - \partial F^{(k)}(\boldsymbol{u}^{(k-1)})$
        $\alpha^0 = \frac{\boldsymbol{p}^T\boldsymbol{p}}{\boldsymbol{p}^T\boldsymbol{y}}$
while $F^{(k)}(u^{(k)} - \alpha^{(k)}\partial F^{(k)}(\boldsymbol{u}^{(k)})) > C^{(k)} - \sigma\alpha^{(k)}(\partial F^{(k)}(\boldsymbol{u}^{(k)}))^T\partial F^{(k)}(\boldsymbol{u}^{(k)})$
        $\alpha^{(k+1)} = \rho\alpha^{(k)}$
end
$\boldsymbol{u}^{(k+1)} = \boldsymbol{u}^{(k)} - \alpha^{(k+1)}\partial F^{(k)}(\boldsymbol{u}^{(k)})$
output $\boldsymbol{u}^{(k+1)}$,
where $F^{(k)}(\boldsymbol{u}) = \frac{\mu}{2}\|\boldsymbol{g} - \boldsymbol{H}\boldsymbol{u}\|_2^2 + \frac{\lambda}{2}\|\boldsymbol{d}^{(k)} - \boldsymbol{B}\nabla\boldsymbol{u} - \boldsymbol{b}^{(k)}\|_2^2$
      $\partial F^{(k)}(\boldsymbol{u}) = \mu\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{u} - \boldsymbol{g}) - \lambda\nabla^T\boldsymbol{B}^T(\boldsymbol{d}^{(k)} - \boldsymbol{B}\nabla\boldsymbol{u} - \boldsymbol{b}^{(k)})$
```

Algorithm 2 (Post-reconstruction PVC with non-smooth optimized PLS regularization)

```
Inputs: PET image $\boldsymbol{g}$, MR image $\boldsymbol{v}$, system psf $\boldsymbol{h}$, $\mu$, $\lambda$
Initialize:
        $\boldsymbol{d}^0 = \boldsymbol{0}$, $\boldsymbol{b}^0 = \boldsymbol{0}$, $\boldsymbol{u}^0 = \boldsymbol{0}$, $\boldsymbol{u}^1 = \boldsymbol{g}$
        $z^0 = 1$, $C^0 = F^0(\boldsymbol{g})$, $\gamma = 0.995$
for k=1:K
    obtain $\boldsymbol{u}^{(k+1)}$ by solving $\boldsymbol{u}$ subproblem with algorithm 1
    $\boldsymbol{d}_{x,y,z}^{(k+1)} = \max\left(\|\boldsymbol{B}\nabla\boldsymbol{u}^{(k)} + \boldsymbol{b}^{(k)}\|_2 - \frac{1}{\lambda}, 0\right)\frac{(\boldsymbol{B}\nabla\boldsymbol{u}^{(k)}+\boldsymbol{b}^{(k)})_{x,y,z}}{\|\boldsymbol{B}\nabla\boldsymbol{u}^{(k)}+\boldsymbol{b}^{(k)}\|_2}$
    $\boldsymbol{b}^{(k+1)} = \boldsymbol{b}^{(k)} + \boldsymbol{B}\nabla\boldsymbol{u}^{(k+1)} - \boldsymbol{d}^{(k+1)}$
    update parameters for algorithm 1:
    $z^{(k+1)} = \gamma z^{(k)} + 1$
    $C^{(k+1)} = \frac{\gamma z^{(k)} C^{(k)} + F^{(k)}(\boldsymbol{u}^{(k+1)})}{z^{(k+1)}}$
    update $\lambda$ with residual balancing
end
```

5.3 Experiments

5.3.1 Simulation Experiments

We first evaluated the proposed method using simulation experiments. We generated a fluorodeoxyglucose (FDG) PET simulation phantom from the BrainWeb phantom [256], as shown in Fig. 5-1 (first row). The phantom was discretized into $181 \times 217 \times 181$ voxels with voxel size of 1 mm × 1 mm × 1 mm. Realistic PET activity distribution was assigned as 12500 Bq/ml in gray matter (GM), 3125 Bq/ml in white matter (WM), 0 Bq/ml in air, CSF and bone, and 1000 Bq/ml in all other tissues [257]. To avoid oversimplified piecewise constant true activity distributions, the activity distribution in GM was further convolved with a Gaussian function with full width at half maximum (FWHM) of 10 mm to create true nonuniformities. Activity in GM was then multiplied with a constant to scale the mean activity back to 12500 Bq/ml. The simulated PET

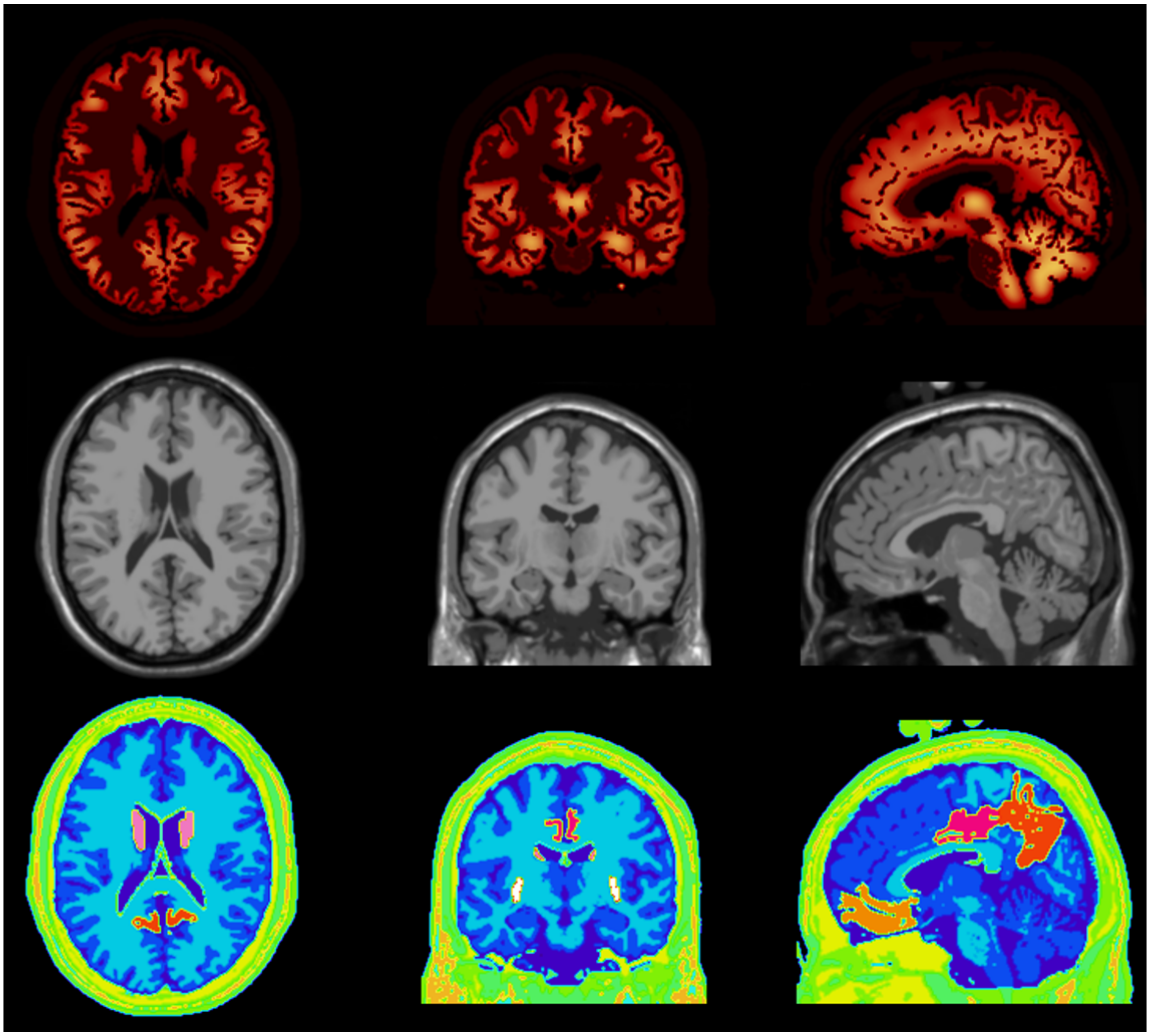


Figure 5-1. Transaxial slice (left column), coronal slice (middle column) and sagittal slice (right column) of simulated PET phantom (top row), MR phantom (middle row) and segmented tissue map (bottom row).

phantom is shown in Fig. 5-1. The simulated phantom was first projected into the sinogram space using a realistic system matrix. Attenuation, scatter and normalization were also modeled within the sinogram space. An MR derived attenuation map was used to simulate the attenuation effect [258] and 25% scatter counts were added. The system resolution was modeled corresponding to a Gaussian PSF with 4.5 mm FWHM. Poisson noise was simulated corresponding to 10^8 total coincidence events. We simulated 20 noise realizations to evaluate quantitative performance of different PVC methods. Reconstructions were performed using ordered-subset expectation maximization (OSEM) algorithm with 10 subsets and 18 iterations.

We compared our proposed methods, which we refer to as 'non-smooth PLS', with: (i) rVC deconvolution, (ii) RBV, (iii) a deconvolution method with asymmetric Bowsher regularization which we refer to as 'aBowsher' method, and (iv) PLS regularized deconvolution using smoothing factor $\beta \neq 0$, which we refer to as 'smooth PLS'. The implementation of rVC followed descriptions in [244], where the step size was chosen as 1.5. The implementation of RBV was same as described in [243]. For 'aBowsher' method, we replace the regularization term in (5.11) with Bowsher type regularization, as described in [259]. In this method, we chosen the search window of size $3 \times 3 \times 3$ and 7 closest gray-level neighbors were picked. In the smooth PLS method, we set smoothing factor $\beta = 10^{-2}$. For the last two methods, the objective function was minimized with limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm [75]. Since rVC and RBV do not apply any smoothing, the reconstructed images were first smoothed using a Gaussian filter with FWHM of 2.5 mm in all three dimensions to suppress noise before being input into these two methods.

Apart from the rVC method, all other methods mentioned above, including our method, requires anatomical guidance. MR image as shown in the middle row of Fig. 5-1 was used as anatomical guidance for our method and the aBowsher method. RBV method requires segmented MR tissue map. Instead of using segmentation map

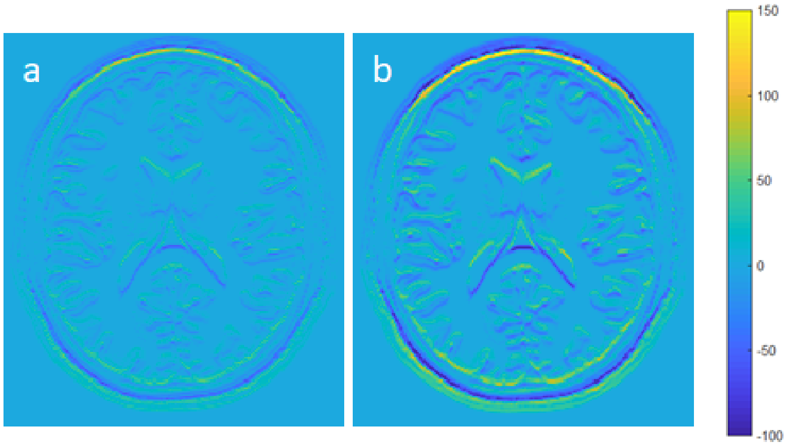


Figure 5-2. Transaxial slice of difference image between perfect registered MR image and misregistered MR image with (a) 1mm and (b) 2mm registration error.

provided by BrainWeb that only contains the entire GM and WM, we segmented the MR image using FreesSurfer to obtain more region of interests (ROIs) [260]. Specifically, we assessed the performance of different PVC methods in medial frontal, precuneus, posterior cingulate, caudate and putamen regions. The bottom row in Fig. 5-1 shows the segmented tissue map provided to RBV and used to study quantitative performance in different ROIs.

In the clinical scenario, one concern for the use of PVC is that it is sensitive to accuracy of MR registration and segmentation [261, 262]. Therefore, in our simulation experiments, we also studied the influence of inaccurate MR registrations and segmentations on the performance of PVC algorithms. We simulated registration mismatch by shifting MR phantom by 1 mm (1 voxel) and 2 mm (2 voxel). The difference images between perfect registered MR image and misregistered MR image are shown in Fig. 5-2. For segmentation error, in order to control the extent of segmentation mismatch, we modeled imperfect MR segmentation with image-wise morphological operations. Specifically, we simulated three different levels of segmentation errors by applying image dilation operation on GM using spheres with radii of 0.2 mm, 0.5 mm and 1 mm, as shown in Fig. 5-3.

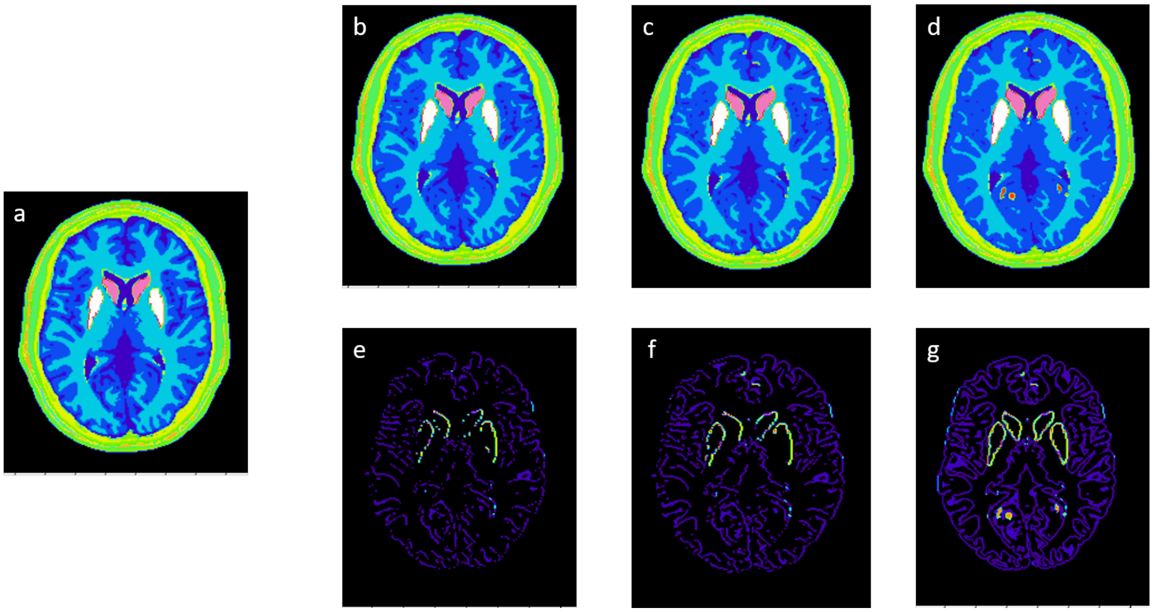


Figure 5-3. Transaxial slice of (a) tissue map with perfect segmentation; (b) tissue map dilated with 0.2 mm sphere; (c) tissue map dilated with 0.5 mm sphere; (d) tissue map dilated with 1 mm sphere; (e) difference image between (a) and (b); (f) difference image between (a) and (c); (g) difference image between (a) and (d).

In order to compare different PVC methods quantitatively, we plotted bias-noise trade-off curves for different ROIs. We computed mean percentage ROI bias as [232]:

$$\text{bias} = \frac{1}{R\theta_{\text{true}}} \sum_{j=1}^{R} (\theta_j - \theta_{\text{true}}) \times 100\%, \tag{5.27}$$

where R is the number of noise realizations, $\theta = \frac{1}{N}\sum_{i=1}^{N} u_i$ denotes mean activity in ROI (N is the number of voxels in the ROI), and θ_{true} is ROI mean activity of ground truth. For noise, we computed average voxel-wise coefficient of variability (cov) [257, 263]:

$$\text{cov} = \frac{1}{\theta_{\text{true}}} \left(\frac{1}{N(R-1)} \sum_{i=1}^{N} \sum_{j=1}^{R} (u_{i,j} - \bar{u}_i)^2 \right)^{\frac{1}{2}} \times 100\%, \tag{5.28}$$

where $\bar{u}_i = \frac{1}{R}\sum_{j=1}^{R} u_{i,j}$ is mean value for voxel i.

5.3.2 Application to *In vivo* Human Dataset

Our assessment of the proposed method on *in vivo* human dataset consists of two parts.

In the first part, we performed a preliminary assessment of our method in few *in vivo* examples. In this part, our data set consists of brain PET scans collected for one 85-year-old cognitively normal amyloid negative male and one 85-year-old cognitively normal amyloid positive female participant in the Baltimore Longitudinal Study of Aging (BLSA). Two scans were performed for each participant, each with a different radiotracer. In the first scan, participants were injected with 370 MBq of ^{15}O-H_2O and data were acquired over 60s once the counts reached threshold levels. In the second scan, 555 MBq of ^{11}C-PIB was injected and data were acquired over 70 min. We use the average of motion-corrected time frames corresponding to 50-70 min post PIB injection in this analysis. The scans were performed on a GE Advance scanner [264] and reconstructed using filtered back projection (FBP) algorithm to generate 3D images of $128 \times 128 \times 35$ voxels with voxel size of 2 mm $\times$ 2 mm $\times$ 4.25 mm. Each participant also underwent a magnetization-prepared rapid gradient echo (MPRAGE) scan on a 3T Philips Achieva scanner (repetition time = 6.8ms, echo time = 3.2ms, flip angle = 8°, image matrix = $256 \times 256 \times 170$, voxel size = $1 \times 1 \times 1 \times 1.2\text{mm}^3$) MPRAGE scans were inhomogeneity-corrected, skull-stripped, and anatomically-labeled using the MUSE algorithm [265]. Each PET scans was rigidly registered onto their corresponding MPRAGE. The inverse of this transform was applied to both the inhomogeneity-corrected (but not skull-stripped) MPRAGE and the anatomical label image to bring them into PET native space. We assessed our non-smooth PLS method, together with RBV and rVC, using these six PET scans. The FWHM of Gaussian PSF was set as 7.5 mm, consistent with measurements on GE Advance scanner [266]. We evaluated visual performances and compare standard uptake value ratios (SUVRs) in different ROIs for these methods.

In the second part, we built a dataset using ^{15}O-H_2O brain PET data from 175 participants from the BLSA. The three PVC methods were further assessed based on the greatest negative age association at baseline (after adjusting for sex) that is given

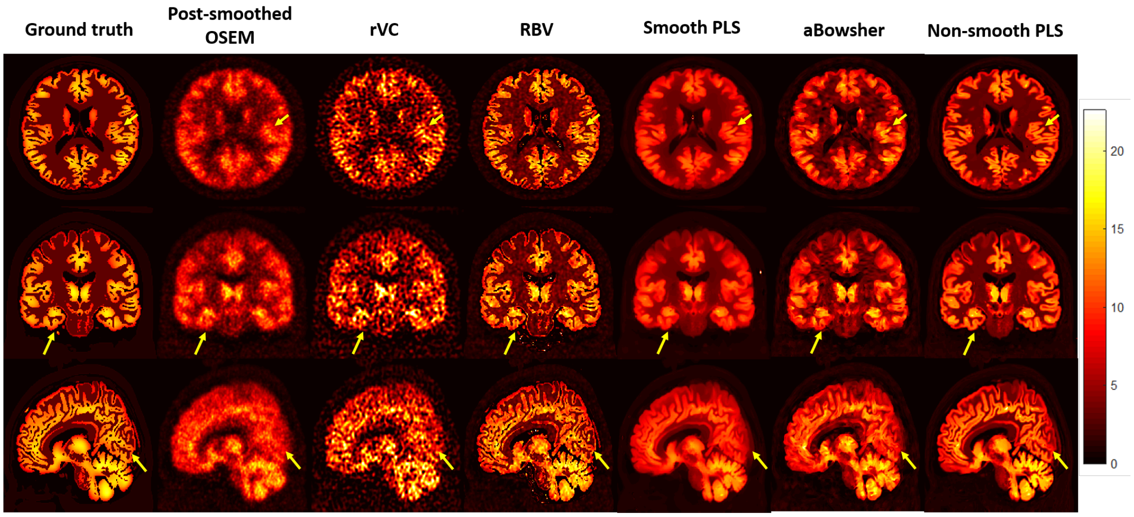


Figure 5-4. Transaxial slice (top row), coronal slice (middle row) and sagittal slice (bottom row) of ground truth (1^{st} column), post-smoothed OSEM (2.5 mm Gaussian filter, 2^{nd} column), rVC (30 iterations, 3^{rd} column), RBV (4^{th} column), smooth PLS ($\mu = 2$, 5^{th} column), aBowsher ($\mu = 2$, 6^{th} column), and non-smooth PLS ($\mu = 60$, 7^{th} column) with perfect MR guidance. Arrows indicate example locations where different structural details can be observed for different methods.

by the number of voxels with statistically significant differences within the cortical gray matter. To do this, voxelwise linear regressions were performed with SUVR as the outcome variable, and age and sex as the independent variables. Multiple comparison correction was performed with 1000 permutation tests. Statistical significance was set when corrected p-value was less than 0.05 for two-sided t-tests.

5.4 Results

5.4.1 Results for Simulation Experiments

Figrue 5-4 shows transaxial, coronal, and sagittal slices for different PVC methods, where perfect MR guidance (no registration or segmentation errors) was used for MR-guided PVC methods. Regularization parameters in smooth/non-smooth PLS and aBowsher were selected so as to balance noise and recovery accuracy in these methods. Compared to post-smoothed OSEM and rVC methods, all MR-guided

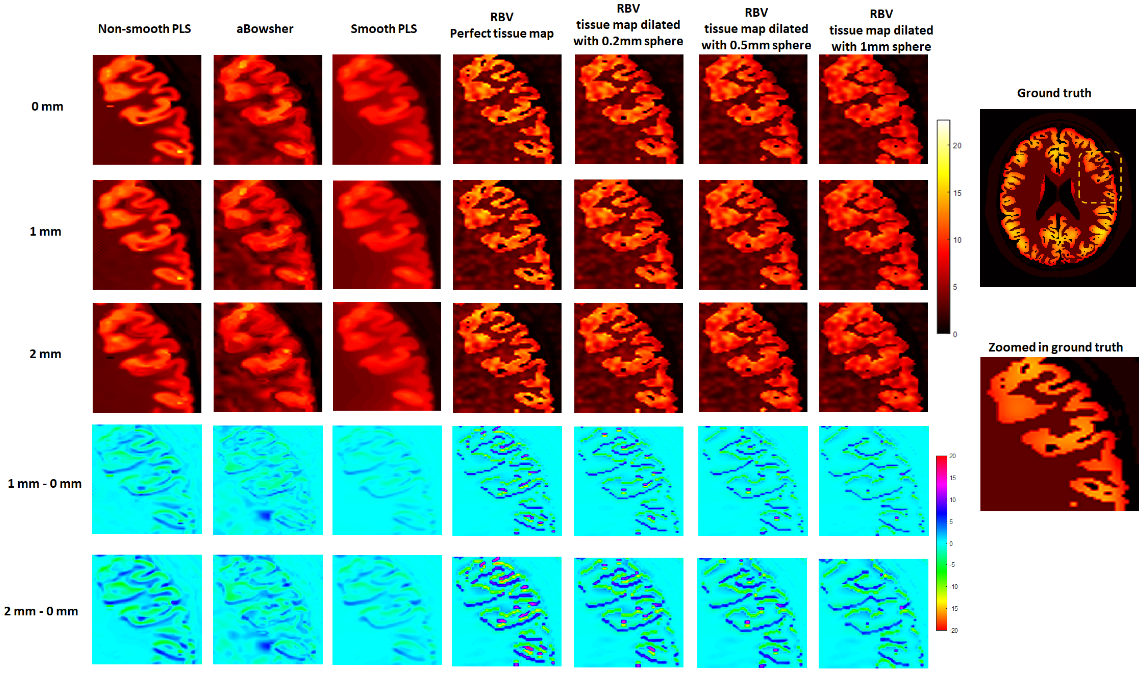

Non-smooth PLS
aBowsher
Smooth PLS
RBV
Perfect tissue map
RBV
tissue map dilated
with 0.2mm sphere
RBV
tissue map dilated
with 0.5mm sphere
RBV
tissue map dilated
with 1mm sphere
0 mm
1 mm
2 mm
1 mm - 0 mm
2 mm - 0 mm
20
15
10
5
0
20
15
10
5
0
-5
-10
-15
-20
Ground truth
Zoomed in ground truth

Figure 5-5. Image patches for different MR-guided PVC methods with different level of registration and segmentation mismatch. First row: MR guidance with 0 mm registration mismatch; Second row: MR guidance with 1 mm registration mismatch; Third row: MR guidance with 2 mm registration mismatch; Fourth row: difference image between first row and second row; Fifth row: difference image between first row and third row.

PVC methods recovered more structure details. Among the four MR-guided PVC methods (smooth/non-smooth PLS, aBowsher and RBV), our proposed non-smooth PLS generated images with comparable structural detail to those obtained using RBV, which requires segmentation of MR image. Compared to the other segmentation-free methods aBowsher and smooth PLS, the proposed method generated images with sharper edges.

Fig. 5-5 illustrates the impact of registration and segmentation errors on different MR-guided PVC methods. In this figure, we quantitatively compare the results of smooth/non-smooth PLS, aBowsher, and RBV using different segmentation maps. The first three rows show zoomed-in image patches of different MR-guided PVC methods with 0 mm, 1 mm and 2 mm registration mismatch, respectively. We then subtracted recovered images with 0 mm registration-mismatched MR guidance from images with 1 mm and 2 mm registration-mismatched MR guidance. These difference images are shown in the 4th and 5th rows. For the 3 segmentation-free methods, edges are seen to become blurry as registration mismatch increases. For RBV, one does not observe blurry edges when there are registration or segmentation errors, but from the difference image one can see that the whole image shifted.

Fig. 5-6 plots bias-noise curves of different PVC methods in different ROIs, with both perfect MR guidance and imperfect MR guidance. For smooth/non-smooth PLS and aBowsher methods, the curves were generated by changing regularization parameter μ. Specifically, we use values in the range 10-500 for non-smooth PLS and 0.6-20 for both smooth PLS and aBowsher methods. For rVC, the curves were generated by changing the number of iterations, from 10 to 60. Note that plots for all three registration mismatch levels are shown with the same range of bias and noise in each ROI in order to reveal the impact of registration mismatch on quantitative performance of different methods. We notice that RBV with perfect tissue map depicts the best bias-noise performance. Imperfect tissue map degrades the performance of RBV. Non-smooth PLS method is inferior to RBV with perfect tissue map but shows better performance compared to other PVC methods. As registration mismatch increases, performance of all MR-guided PVC methods degrades and differences between these methods become less obvious.

5.4.2 Results for In vivo Human Dataset

Fig. 5-7 shows transaxial, coronal and sagittal slices for the ^{15}O-H$_2$O scans with different PVC methods. Fig. 5-8 shows similar results for ^{11}C-PIB. For both ^{15}O-H$_2$O and ^{11}C-PIB scans, we observed our proposed non-smooth PLS to generate images with lower noise. For example, if we quantify noise using standard deviation in postcentral gyrus, for ^{15}O-H$_2$O scans, non-smooth PLS generated images with spatial noise on average 49% lower compared to RBV method, and 69% lower compared to

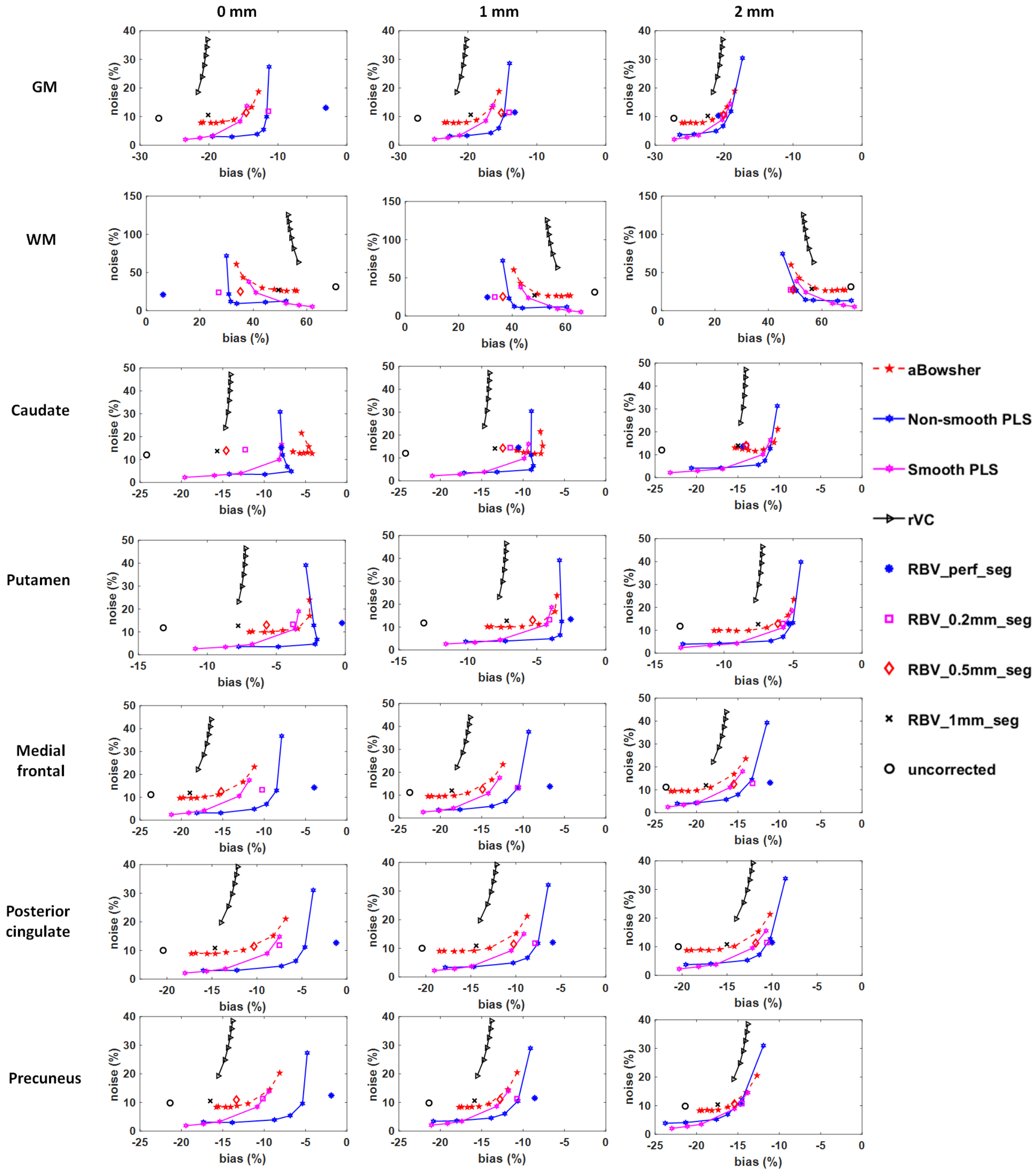


Figure 5-6. Bias-noise (cov) trade-off curves for different PVC methods in different ROIs with MR registration mismatch of 0 mm (first column), 1 mm (second column) and 2 mm (third column).

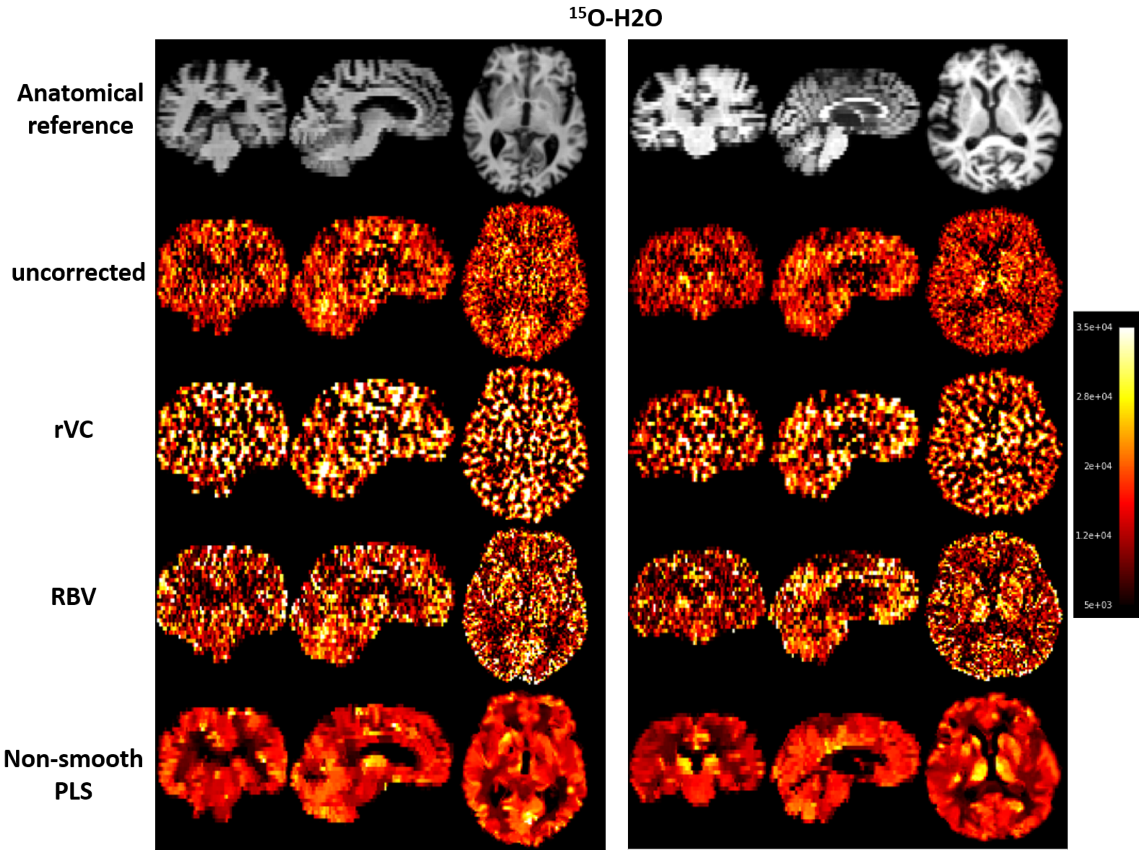

15O-H2O
Anatomical reference
uncorrected
rVC
RBV
Non-smooth PLS
3.5e+04
2.8e+04
2e+04
1.2e+04
5e+03

Figure 5-7. PVC results on $^{15}\mathrm{O}\text{-}\mathrm{H_2O}$ scans for an amyloid negative (left) and amyloid positive (right) participant.

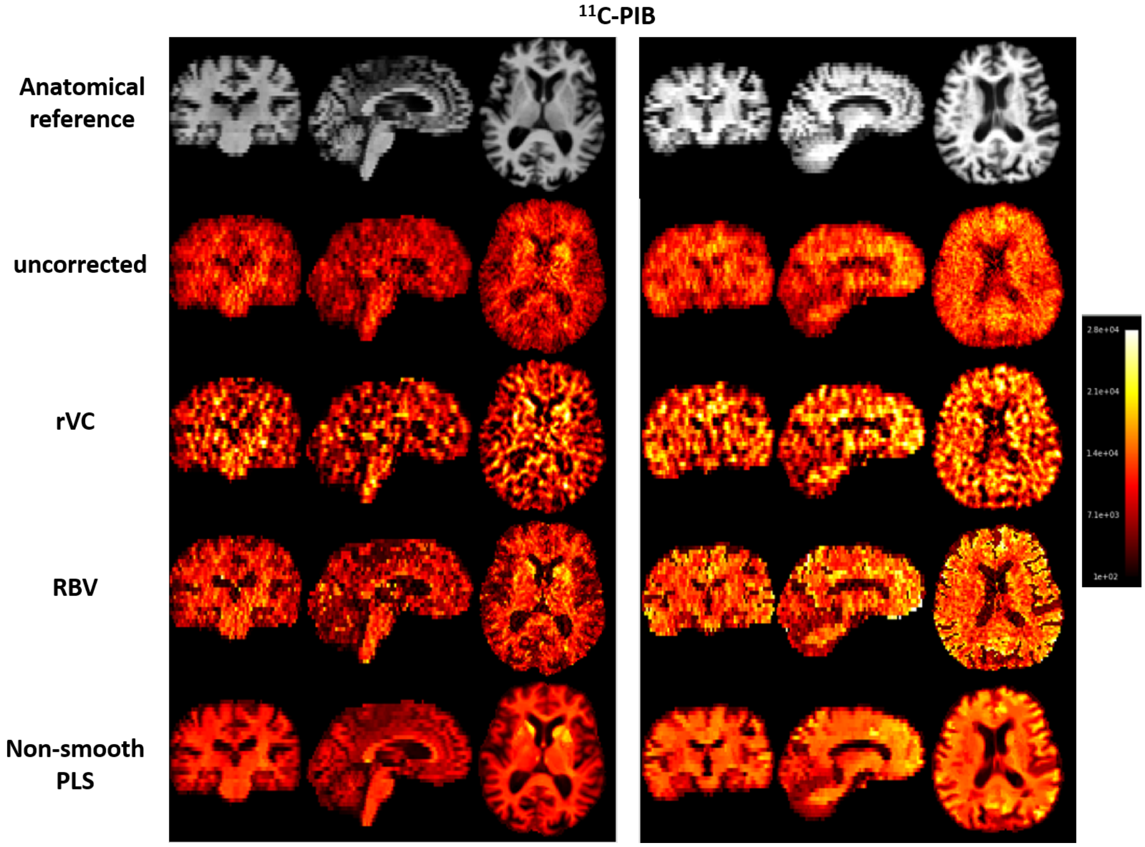
RVC method. For ^{11}C-PIB scans, images from non-smooth PLS have on average 27% lower noise compared to RBV method and 59% lower noise compared to RVC method. Meanwhile, non-smooth PLS shows comparable ROI mean activities to those of RBV and RVC method. For example, in the posterior cingulate region, mean activity from non-smooth PLS is on average 0.7% higher than RVC method and 13% lower than RBV method for ^{15}O-H$_2$O scans, and is on average 0.3% higher than RVC method and 9% lower than RBV method for ^{11}C-PIB scans. In the caudate region, mean activity from non-smooth PLS is on average 2% higher than RVC method and 5% lower than RBV method for ^{15}O-H$_2$O scans, and is on average 2% lower than RVC method and 7% lower than RBV method for ^{11}C-PIB scans.

Fig. 5-9 plots SUVRs from different PVC methods for amyloid negative and amyloid positive participants and their differences. It can be observed that different PVC methods lead to increased SUVR compared to uncorrected images. Our method, together with other PVC methods, provide larger differences between amyloid positive and amyloid negative scans compared to uncorrected image.

Fig. 5-10 shows the normalized distribution of two-sided t-values for the association between age and SUVR, and the spatial distribution of statistically significant voxels for different PVC methods.

5.5 Discussion

In this work, we developed a post-reconstruction PVC method based on deconvolution with PLS regularization. The method avoids assumption on uniform activity distribution and requirement of MR segmentation, which are used by most current post-reconstruction PVC methods. Further, we developed a non-smooth optimization algorithm based on split Bregman method so that the method could be efficiently applied on 3D image. Our proposed method was evaluated using both simulated and



¹¹C-PIB
Anatomical reference
uncorrected
rVC
RBV
Non-smooth PLS
2.8e+04
2.1e+04
1.4e+04
7.1e+03
1e+02

Figure 5-8. PVC results on ^{11}C-PIB scans for an amyloid negative (left) and amyloid positive (right) participant.

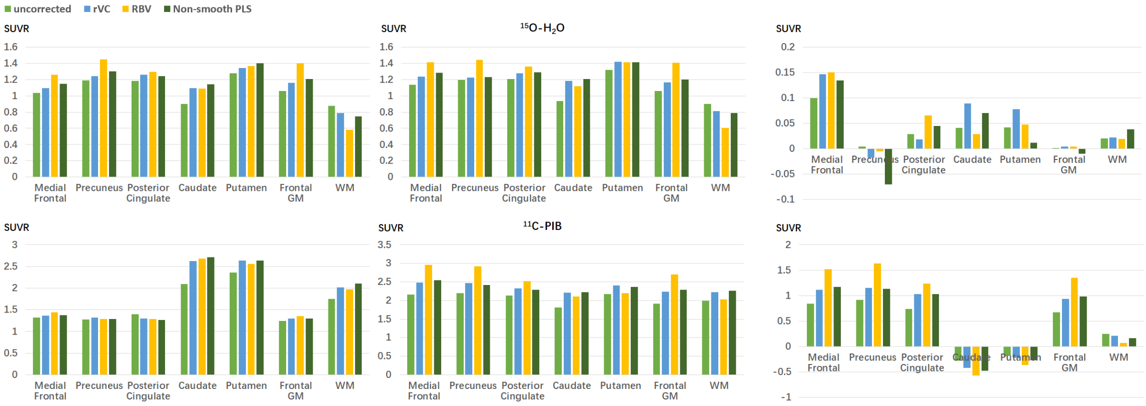


Figure 5-9. Mean concentration in different ROIs for different PVC methods for amyloid negative (left), amyloid positive (middle) participants and their difference (right).

real data.

From our simulation experiment results, we first observe that the use of anatomical information from MRI brings substantial improvement both visually and quantitatively, as shown in Fig. 5-4 and Fig. 5-6. Specifically, although rVC method, which does not use MR guidance, leads to improved image contrast and smaller bias compared to uncorrected images, the images obtained with this method are significantly noisier. In contrast, the four MR-guided PVC methods are able to control spatial noise amplifications while further improve quantitative accuracy. Even when reasonable registration errors occurred in MR guidance, the improvement brought about by MR information was still considerable.

We next compare the four MR-guided methods: smooth/non-smooth PLS, aBowsher and RBV. As mentioned previously, when perfect MR guidance is provided, RBV method shows best performance, while our proposed non-smooth PLS method performs better than the other methods. Specifically, if we compare non-smooth PLS with smooth PLS, we clearly see the improvement brought about by non-smooth optimization algorithm: both sharper edges in visual results as well as better bias-noise trade-off performance in quantitative results. For aBowsher method, overall its performance is better than smooth PLS but inferior to non-smooth PLS method. It shows the best quantitative performance in the caudate region, but is outperformed by non-smooth PLS and RBV with perfect tissue map in other regions.

In real scenario, perfect MR guidance is never available. We also studied impact of imperfect MR guidance on different MR-guided PVC methods. As mentioned previously, it causes additional blur in output images for segmentation-free PVC methods, and shifted images for RBV method. Quantitatively, imperfection from registration and segmentation mainly degrades bias of MR-guided PVC methods. From Fig. 5-6, we observe that although RBV with perfect tissue map shows the best performance, at 2 mm registration mismatch, its performance almost overlaps with that of non-smooth PLS in several ROIs, such as in WM, putamen, posterior cingulate and precuneus. The mismatch in segmentation also degrades the performance of RBV. For tissue map dilated with 0.2 mm sphere, the performance of RBV is similar to that of non-smooth PLS in several situations, such as in GM in the presence of a 0- or 1-mm registration error, and in medial frontal, posterior cingulate and precuneus in the presence of a 1- or 2-mm registration error. In ROIs other than WM, RBV shows inferior performance to that of non-smooth PLS. As segmentation mismatch further increases, with tissue map dilated with 0.5 mm sphere, performance of RBV is similar to aBowsher method in several ROIs, including the GM, putamen, medial frontal, posterior cingulate and precuneus. For RBV with tissue map dilated using 1 mm sphere, though it still shows improvement compared to post-smoothed OSEM, its improvement on bias is similar as rVC in most cases, which does not use MR guidance, regardless of the registration mismatch level. These indicate that the performance of segmentation-based PVC method, such as RBV, is sensitive to segmentation accuracy. On the other hand, segmentation-free methods such as smooth/non-smooth PLS and aBowsher, as explored in this work, are not susceptible to segmentation errors.

Our method was further evaluated using *in vivo* human data. Our non-smooth PLS PVC method yields images with better visual quality, which retain the structural detail while control the image noise at lower level, as shown in Fig. 5-7 and 5-8. Our method also shows potential to provide larger difference between amyloid positive and amyloid negative scans comparing to images without PVC, as shown in Fig. 5-9 Furthermore, from Fig. 5-10, we observed the non-smooth PLS method yields the greatest number of statistically significant voxels where older individuals had lower SUVR. This indicates our method may enable discovery of effects that would not be considered statistically significant when no PVC or other PVC methods are applied.

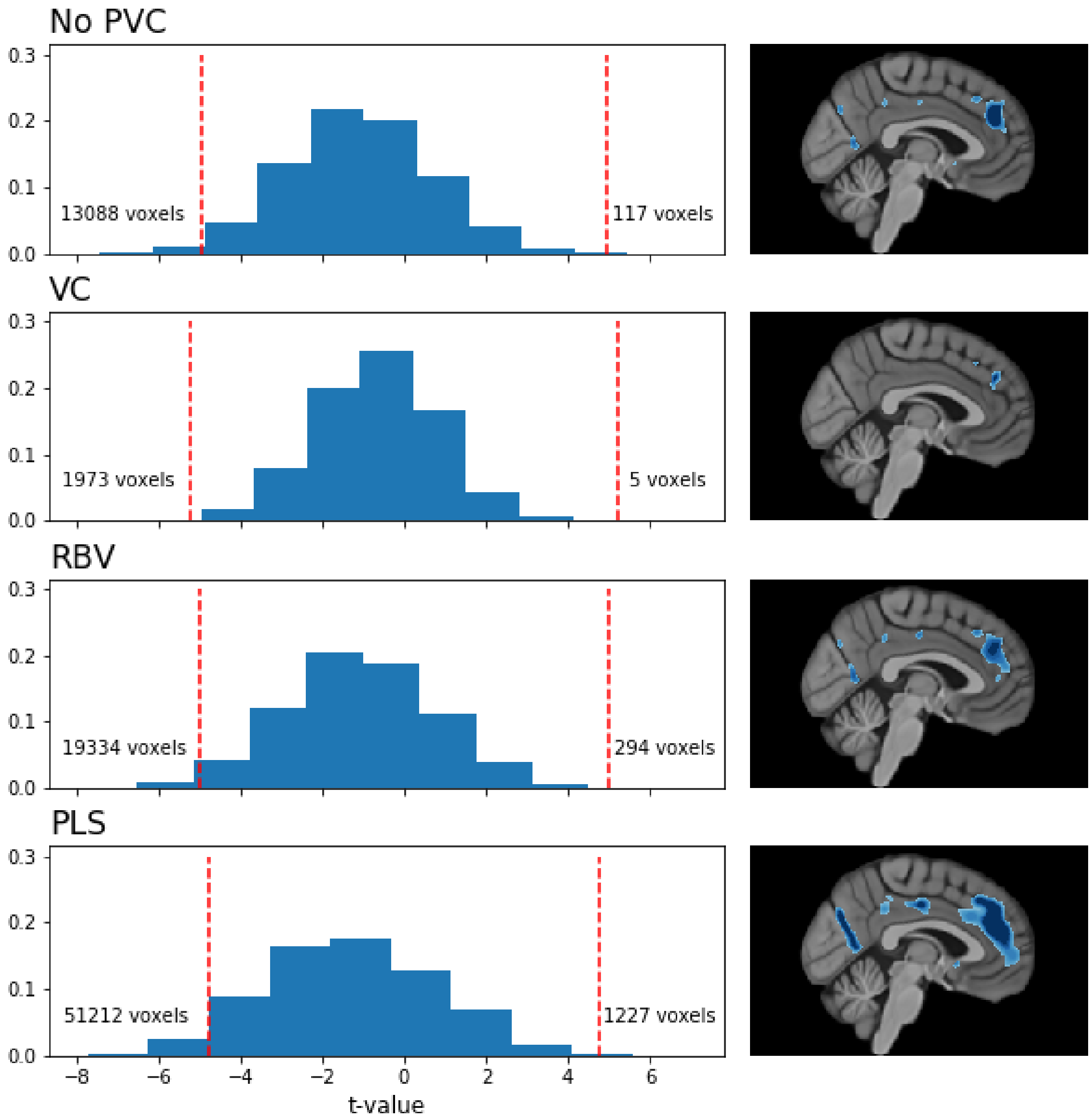


Figure 5-10. Normalized histograms of two-sided t-values for the age term in the linear regression model with voxelwise SUVR as the outcome (adjusted for sex). Each row corresponds to a different PVC method. Red vertical dashed lines indicate the t-values that correspond to a multiple comparison-corrected (via permutation tests) $p < 0.05$ threshold. Number of statistically significant cortical gray matter voxels with a negative and positive association are indicated to the left and right of the dashed lines, respectively. Sagittal slices illustrate the extent of the statistically significant clusters on the left medial cortical surface.

5.6 Conclusion

We have presented a deconvolution-based post-reconstruction PVC method with PLS regularization. A non-smooth optimization method based on split Bregman algorithm is used to effectively solve the regularized deconvolution problem as well as avoiding additional blur caused by smoothing. We thoroughly quantified the performance of our method in the case of both perfect and imperfect MR guidance, and compared it against other voxel-wise PVC methods. Results from simulation experiments indicate our method has promise to show advanced performance when realistic anatomical guidance is provided. The proposed method was also evaluated with real data and show promise performance in the analysis of scans acquired with different radiotracers. Our results based on simulated and real data provide further confidence regarding applicability of our methods to routine clinical PET imaging.

Chapter 6

Conclusions and general discussion

6.1 Chapter 3: Incorporating boundary condition for Neumann-series-based RTE

We have proposed a Neumann-series-based RTE with boundary condition modeled to improve photon propagation computations in tissue. In addition, we developed an algorithm for reconstruction of optical properties based on the photon propagation model we developed. The proposed method was estimated using a Monte Carlo simulated 3D DOI setup. The results indicate the proposed method yield more accurate photon propagation modeling compared to Neumann-series-based RTE without boundary condition. Futhermore, in the task of reconstructing absorption coefficients for the 3D DOI system, the proposed method generates substantially more accurate estimation.

A limitation in our work is that the method was only implemented in homogeneous mediums, while in practice different biophotonics applications require modeling photon propagation in heterogeneous mediums. One future work is to extend this model for heterogeneous medium. Here we describe one possible approach to this. Consider a heterogeneous medium with K sub-domains, where each sub-domain is characterized by a different refractive index. Similar to the method in [189], we can derive a set of

K coupled equations as

$$\begin{cases} w_1 = \boldsymbol{\mathcal{X}}\Xi_1 + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}w_1 + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}w_1 \\ w_2 = \boldsymbol{\mathcal{X}}\Xi_2 + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}w_2 + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}w_2 \\ \vdots \\ w_k = \boldsymbol{\mathcal{X}}\Xi_k + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{K}}w_k + \boldsymbol{\mathcal{X}}\boldsymbol{\mathcal{R}}w_k, \end{cases} \tag{6.1}$$

where w_i is distribution function in i^{th} tissue, Ξ_i is the source term in i^{th} tissue, which can be written as $\Xi_i = S_i + \boldsymbol{T}w_{icb}$, where S_i is any direct contribution due to the laser source in tissue i, and w_{icb} is distribution function at boundary from any neighboring tissue, and $\boldsymbol{T}$ is transmission operator. Solving this set of coupled equations will provide distributions in heterogeneous medium. Further, current Neumann-series framework may require large memory and long computation times when processing mediums with high scattering coefficients or large geometry. In these cases, diffuse-approximation-based methods could generate accurate results. Another research direction is to integrate the Neumann-series method with diffuse approximation methods. In addition, although Neumann-series-based RTE without boundary condition has been implemented using GPU [55], the model we developed in this work currently only has CPU-based implementation. Extending the implementation to GPU could reduce computation time and potentially allow this model to be applied in more complex scenarios. This is also an important future work.

6.2 Chapter 4: Reconstruction algorithm for FMT

In this chapter, we developed a reconstruction algorithm for FMT based on sparsity-initialized MLEM. We evaluated our proposed method together with uniform-initialized MLEM method and pure sparse recovery method using both homogeneous and heterogeneous simulation phantoms. Apart from improved qualitative and quantitative performance, the proposed method shows several other advantages. Compared to uniform-initialized MLEM method, the proposed method is faster to execute. Compared to pure sparse reconstruction, the method is more robust to noise amplification caused by preconditioning. In addition, our study also demonstrates the importance of accurately modeling Poisson noise in FMT applications.

In our work, we assumed that there is no background fluorescence signal. However, in real FMT applications, it is possible that weak background fluorescence signals exist. Generalizing our current method to the case with background fluorescence is a future research direction. When background fluorescence occurs, our existing assumption on spatial sparsity might become invalid. Possible alternatives include exploring sparsity in other domains such as in the gradient space or wavelet space. Furthermore, we evaluated the proposed method with extensive simulation studies. Evaluating the performance of the proposed method using physical phantom and *in vivo* animal experiments is another important future work. Finally, in this work, we obtained the forward model using a Monte-Carlo-based framework. Combining the proposed reconstruction method with analytical-based forward model such as the Neumann-series RTE which we described in Chapter 3 is another direction for future research.

6.3 Chapter 5: Partial volume correction for brain PET imaging

We developed a deconvolution-based post-reconstruction PVC method with PLS regularization. Both simulation and *in vivo* human data were used to evaluate our proposed method, together with several other voxel-wise PVC methods. Results demonstrate the proposed method provides fine structure detail while reducing image noise. In addition, the proposed method shows resilience to mismatches between anatomical and functional images. Last, voxel-wise statistical analysis indicated the proposed method to have great potential to reveal voxel-wise effects that would not be considered statistically significant with other PVC methods.

Future work involves evaluating the proposed PVC method more thoroughly on clinical datasets. In this work, we only compared regional standard uptake value ratio of different PVC methods on a very limited dataset. In future, we plan to evaluate the mean and variability of regional standard uptake value ratio on larger dataset. Furthermore, we plan to apply and extend a so-called no-gold-standard technique [267, 268] to evaluate quantitative performance such as bias and noise of different PVC methods on clinical dataset. More regional and voxel-wise statistical analysis will also be performed to study if our proposed PVC method improves separability between clinical participants with dementia disease from those with normal aging. In addition, deep learning has attracted much research interest for its wide applications in PET [269–272]. Recently, deep-learning-based method has been developed to generate super-resolution PET image from low-resolution image [273, 274]. These methods show promising results and provde alternative ways to perform partial volume correction. Comparing our proposed method with deep-learning-based methods is another future research.

Appendix A

Application of computational breast phantoms to evaluate reconstruction methods for fluorescence molecular tomography

Fluorescence molecular tomography (FMT) has potential of providing high contrast images for breast tumor detection. Computational phantom provides a convenient way to a wide variety of fluorophore distribution configurations in patients and perform comprehensive evaluation of the imaging systems and methods for FMT. In this study, a digital breast phantom was used to compare the performance of a novel sparse reconstruction method and Tikhonov regularization method for resolving tumors with different amount of separation. The results showed that the sparse reconstruction method yielded better performance. This simulation-based approach with computational phantoms enabled an evaluation of the reconstruction methods for FMT for breast-cancer detection.

A.1 Introduction

Fluorescence imaging is an emerging optical technique that provides high contrast image with the use of fluorophores. Previously, Indocyanine Green (ICG) has been reported for fluorescence image-guided surgery for breast cancer [275]. Recently, clinical studies have been conducted to study the feasibility of fluorescence molecular tomography (FMT) for breast tumor detection [27]. The advent of sparse reconstruction methods has shown potential of providing more accurate estimation of the location and intensity of fluorescence distribution compared to conventional method in FMT [83]. However, studies are required to assess whether the sparse reconstruction methods indeed yield improved performance for FMT-based breast-cancer imaging. Computational phantom provides a comparatively easier way to conduct these studies by allowing modeling of different fluorscence distributions in breast-cancer populations as opposed to having to actually image patients. In this study, we conducted a computational human-phantom-based study to compare the performance of sparse reconstruction method and a traditional Tikhonov regularization method for resolving tumors in patient images.

A.2 Method

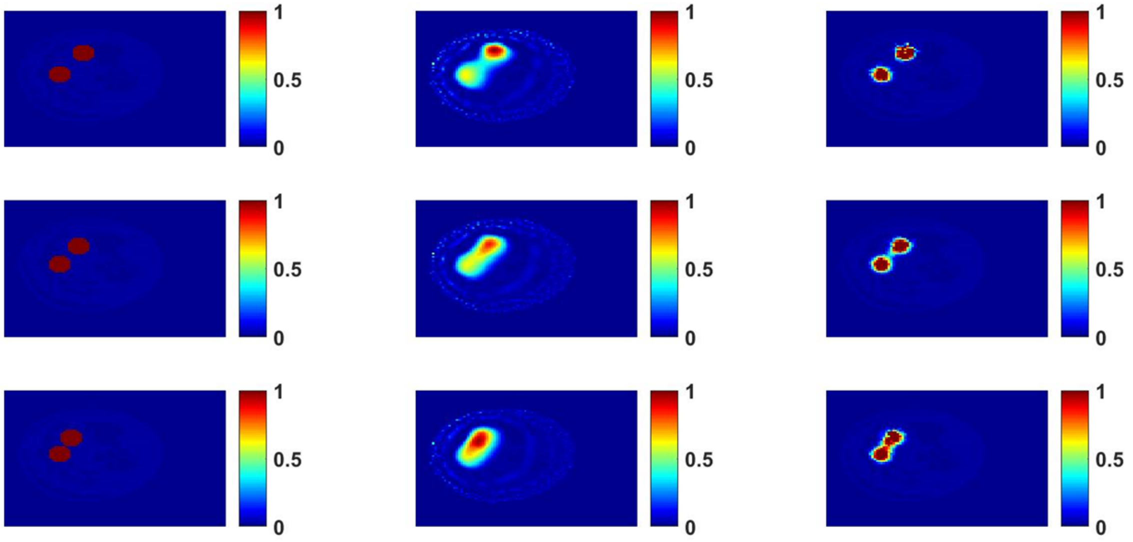
As we have described in Chapter 2, the forward model of FMT is given by

$$\mathbf{\Phi} = \mathbf{G}\boldsymbol{x}, \tag{A.1}$$

where $\mathbf{G}$ is system matrix, $\mathbf{\Phi}$ is detector measurements, and $\boldsymbol{x}$ is unknown fluorescence yield. To reconstruct the fluorescence yield, the following optimization problem with Tikhonov regularization is commonly used:

$$\hat{x} = \arg\min_{x} \|\mathbf{G}\boldsymbol{x} - \mathbf{\Phi}\|_2^2 + +\lambda\|\mathbf{\Gamma}\boldsymbol{x}\|_2^2, \tag{A.2}$$

where $\mathbf{\Gamma}$ is a weighting matrix. In this case, we choose identity matrix $\boldsymbol{I}$ as the weighting matrix. For sparse reconstruction method, we choose the sparsity-initialized MLEM algorithm we described in Chapter 4.



(a) True image

(b) Tikhonov regularized reconstruction

(c) Sparsity-initialized MLEM method

Figure A-1. Comparison of reconstruction with Tikhonov regularization method and the sparse reconstruction method. From top row to bottom row, the distances of the center of two tumors are 2.5 cm, 2 cm and 1.7 cm, respectively.

A.3 Experiments

Simulation-based experiments were conducted to study the ability of FMT for resolving breast tumor. A breast phantom generated from MR database was used in this experiment. The phantom was segmented into four different tissues: blood vessels, skin layer, fat, and fibroglandular tissues. Different optical properties were assigned to each type of tissue, as described in [276]. Two tumors of size 1.6 cm were inserted in the breast phantom at difference separation and fluorophore was assumed to concentrate in the tumors. 40 laser sources and 40 detectors were positioned around the phantom. Monte Carlo method was implemented to calculate the forward model [68]. Two different methods, Tikhonov regularization method and the sparsity-initialized MLEM reconstruction method were used for reconstruction of FMT.

A.4 Results and conclusions

The result is shown in Fig. A-1. First, we observe while Tikhonov regularization generate images with oversmoothed edges, the sparse reconstruction method is able to preserve edges of the two tumors. Further, we notice that for larger separation between tumors (2.5cm), both Tikhonov method and sparse reconstruction method can differentiate the two tumors. When the separation between the tumors decreases, it becomes hard to distinguish the two tumors from images produced by Tikhonov regularization. However, the sparse reconstruction method is still able to resolve the two tumors. In addition, we also notice that compared to Tikhonov reconstruction method, sparse reconstruction provides more accurate estimation of fluorophore distribution and signal intensity.

In this work, we evaluated the performance of our proposed sparse reconstruction method for the task of resolving breast tumor and compared it with commonly used Tikhonov regularization method. Although there have been studies to apply FMT for the task of human breast tumor detection, the use of sparse reconstruction technique is very limited in such applications [27, 70]. The results in this study indicate that sparse reconstruction technique has potential to provide better performance in this task. In future, we will further evaluate if sparse reconstruction methods outperform conventional reconstruction method for the task of breast tumor detection in more realistic circumstance, such as in the case of phantom study and *in vivo* human study.

References

[1] M. L. James and S. S. Gambhir, “A molecular imaging primer: modalities, imaging agents, and applications,” *Physiological reviews*, vol. 92, no. 2, pp. 897–965, 2012.

[2] T. F. Massoud and S. S. Gambhir, “Molecular imaging in living subjects: seeing fundamental biological processes in a new light,” *Genes & development*, vol. 17, no. 5, pp. 545–580, 2003.

[3] S. Vallabhajosula, *Molecular imaging: radiopharmaceuticals for PET and SPECT*. Springer Science & Business Media, 2009.

[4] A. W. Sauter, H. F. Wehrl, A. Kolb, M. S. Judenhofer, and B. J. Pichler, “Combined PET/MRI: one step further in multimodality imaging,” *Trends in molecular medicine*, vol. 16, no. 11, pp. 508–515, 2010.

[5] F. Hyafil, J.-C. Cornily, J. E. Feig, R. Gordon, E. Vucic, V. Amirbekian, E. A. Fisher, V. Fuster, L. J. Feldman, and Z. A. Fayad, “Noninvasive detection of macrophages using a nanoparticulate contrast agent for computed tomography,” *Nature medicine*, vol. 13, no. 5, pp. 636–641, 2007.

[6] K. A. Kelly, J. R. Allport, A. Tsourkas, V. R. Shinde-Patil, L. Josephson, and R. Weissleder, “Detection of vascular adhesion molecule-1 expression using a novel multimodal nanoparticle,” *Circulation research*, vol. 96, no. 3, pp. 327–336, 2005.

[7] N. Deshpande, A. Needles, and J. K. Willmann, “Molecular ultrasound imaging:

current status and future directions," *Clinical radiology*, vol. 65, no. 7, pp. 567–581, 2010.

[8] F. Stuker, J. Ripoll, and M. Rudin, "Fluorescence molecular tomography: principles and potential for pharmaceutical research," *Pharmaceutics*, vol. 3, no. 2, pp. 229–274, 2011.

[9] A. B. Rosenkrantz, M. Mendiratta-Lala, B. J. Bartholmai, D. Ganeshan, R. G. Abramson, K. R. Burton, J. Y. John-Paul, E. M. Scalzetti, T. E. Yankeelov, R. M. Subramaniam *et al.*, "Clinical utility of quantitative imaging," *Academic radiology*, vol. 22, no. 1, pp. 33–49, 2015.

[10] A. K. Jha, Y. Zhu, J. U. Kang, J. K. Dreyer, A. Gjedde, D. F. Wong, and A. Rahmim, "Incorporating boundary conditions in the integral form of the radiative transfer equation for transcranial imaging," in *Clinical and Translational Biophotonics*. Optical Society of America, 2016, pp. JW3A–47.

[11] A. K. Jha, Y. Zhu, D. F. Wong, and A. Rahmim, "A radiative transfer equation-based image-reconstruction method incorporating boundary conditions for diffuse optical imaging," in *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, vol. 10137. International Society for Optics and Photonics, 2017, p. 1013705.

[12] A. K. Jha, Y. Zhu, S. Arridge, D. F. Wong, and A. Rahmim, "Incorporating reflection boundary conditions in the neumann series radiative transport equation: application to photon propagation and reconstruction in diffuse optical imaging," *Biomedical optics express*, vol. 9, no. 4, pp. 1389–1407, 2018.

[13] Y. Zhu, A. K. Jha, J. K. Dreyer, H. N. Le, J. U. Kang, P. E. Roland, D. F. Wong, and A. Rahmim, "A three-step reconstruction method for fluorescence molecular tomography based on compressive sensing," in *Optical Tomography and Spectroscopy of Tissue XII*, vol. 10059. International Society for Optics

and Photonics, 2017, p. 1005911.

[14] Y. Zhu, A. K. Jha, D. F. Wong, and A. Rahmim, “Improved sparse reconstruction for fluorescence molecular tomography with poisson noise modeling,” in *Microscopy Histopathology and Analytics.* Optical Society of America, 2018, pp. JTu3A–51.

[15] ——, “Image reconstruction in fluorescence molecular tomography with sparsity-initialized maximum-likelihood expectation maximization,” *Biomedical optics express*, vol. 9, no. 7, pp. 3106–3121, 2018.

[16] Y. Zhu, Y. Gao, O. Rousset, D. Wong, and A. Rahmim, “Post-reconstruction MRI-guided enhancement of PET images using parallel level set method with Bregman iteration,” *Journal of Nuclear Medicine*, vol. 60, no. supplement 1, pp. 179–179, 2019.

[17] Y. Zhu and A. Rahmim, “MR-guided partial volume correction of 3D PET images using a split bregman optimized parallel level set framework,” in *EUROPEAN JOURNAL OF NUCLEAR MEDICINE AND MOLECULAR IMAGING*, vol. 46, no. SUPPL 1. SPRINGER 233 SPRING ST, NEW YORK, NY 10013 USA, 2019, pp. S788–S788.

[18] Y. Zhu, Y. Gao, and A. Rahmim, “Subtle MR guidance for partial volume correction of PET images: A comparison of techniques,” in *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC).* IEEE, pp. 1–3.

[19] M. Bilgel, Y. Zhu, O. Rousset, D. F. Wong, S. M. Resnick, and A. Rahmim, “Objective evaluation of voxel-wise partial volume correction methods in brain PET imaging,” *Journal of Nuclear Medicine*, vol. 61, no. supplement 1, pp. 72–72, 2020.

[20] Y. Zhu, M. Bilgel, Y. Gao, O. G. Rousset, S. M. Resnick, D. F. Wong, and

A. Rahmim, “Deconvolution-based partial volume correction of PET images with parallel level set regularization,” *IEEE transactions on medical imaging*, (submitted).

[21] Y. Zhu, A. K. Jha, and A. Rahmim, “Application of computational breast phantoms to evaluate reconstruction methods for fluorescence molecular tomography,” *arXiv preprint arXiv:1803.11289*, 2018.

[22] M. O’Leary, D. Boas, B. Chance, and A. Yodh, “Reradiation and imaging of diffuse photon density waves using fluorescent inhomogeneities,” *Journal of Luminescence*, vol. 60, pp. 281–286, 1994.

[23] X. Li, M. O’leary, D. Boas, B. Chance, and A. Yodh, “Fluorescent diffuse photon density waves in homogeneous and heterogeneous turbid media: analytic solutions and applications,” *Applied Optics*, vol. 35, no. 19, pp. 3746–3758, 1996.

[24] M. A. Pysz, S. S. Gambhir, and J. K. Willmann, “Molecular imaging: current status and emerging strategies,” *Clinical radiology*, vol. 65, no. 7, pp. 500–516, 2010.

[25] D. Hyde, R. de Kleine, S. A. MacLaurin, E. Miller, D. H. Brooks, T. Krucker, and V. Ntziachristos, “Hybrid FMT–CT imaging of amyloid-β plaques in a murine Alzheimer’s disease model,” *Neuroimage*, vol. 44, no. 4, pp. 1304–1311, 2009.

[26] S. Raymond, A. Kumar, D. Boas, and B. Bacskai, “Optimal parameters for near infrared fluorescence imaging of amyloid plaques in Alzheimer’s disease mouse models,” *Physics in Medicine & Biology*, vol. 54, no. 20, p. 6201, 2009.

[27] A. Corlu, R. Choe, T. Durduran, M. A. Rosen, M. Schweiger, S. R. Arridge, M. D. Schnall, and A. G. Yodh, “Three-dimensional in vivo fluorescence diffuse optical tomography of breast cancer in humans,” *Optics express*, vol. 15, no. 11, pp. 6696–6716, 2007.

[28] X. Montet, V. Ntziachristos, J. Grimm, and R. Weissleder, "Tomographic fluorescence mapping of tumor targets," *Cancer Research*, vol. 65, no. 14, pp. 6330–6336, 2005.

[29] https://svi.nl/Fluorescence.

[30] J. R. Lakowicz, *Principles of fluorescence spectroscopy.* Springer Science & Business Media, 2013.

[31] N. Deliolanis, T. Lasser, D. Hyde, A. Soubret, J. Ripoll, and V. Ntziachristos, "Free-space fluorescence molecular tomography utilizing 360 geometry projections," *Optics letters*, vol. 32, no. 4, pp. 382–384, 2007.

[32] T. Lasser, A. Soubret, J. Ripoll, and V. Ntziachristos, "Surface reconstruction for free-space 360° fluorescence molecular tomography and the effects of animal motion," *IEEE transactions on medical imaging*, vol. 27, no. 2, pp. 188–194, 2008.

[33] R. Bourayou, H. Boeth, H. Benav, T. Betz, U. Lindauer, T. Nierhaus, J. Klohs, A. Wunder, U. Dirnagl, and J. M. Steinbrink, "Fluorescence tomography technique optimized for noninvasive imaging of the mouse brain," *Journal of Biomedical Optics*, vol. 13, no. 4, p. 041311, 2008.

[34] R. B. Schulz, A. Ale, A. Sarantopoulos, M. Freyer, E. Soehngen, M. Zientkowska, and V. Ntziachristos, "Hybrid system for simultaneous fluorescence and x-ray computed tomography," *IEEE transactions on medical imaging*, vol. 29, no. 2, pp. 465–473, 2009.

[35] S. C. Davis, K. S. Samkoe, J. A. O'Hara, S. L. Gibbs-Strauss, H. L. Payne, P. J. Hoopes, K. D. Paulsen, and B. W. Pogue, "MRI-coupled fluorescence tomography quantifies EGFR activity in brain tumors," *Academic radiology*, vol. 17, no. 3, pp. 271–276, 2010.

[36] Y. Lin, M. Ghijsen, H. Gao, N. Liu, O. Nalcioglu, and G. Gulsen, "Photo-

multiplier tube based hybrid MRI and frequency domain fluorescence tomography system for small animal imaging," *Physics in medicine and biology*, vol. 56, no. 15, p. 4731, 2011.

[37] V. Venugopal, J. Chen, and X. Intes, "Development of an optical imaging platform for functional imaging of small animals using wide-field excitation," *Biomedical optics express*, vol. 1, no. 1, p. 143, 2010.

[38] R. Yao, Q. Pian, and X. Intes, "Wide-field fluorescence molecular tomography with compressive sensing based preconditioning," *Biomedical optics express*, vol. 6, no. 12, pp. 4887–4898, 2015.

[39] F. Leblond, S. C. Davis, P. A. Valdés, and B. W. Pogue, "Pre-clinical whole-body fluorescence imaging: Review of instruments, methods and applications," *Journal of photochemistry and photobiology B: Biology*, vol. 98, no. 1, pp. 77–94, 2010.

[40] A. X. Cong and G. Wang, "A finite-element-based reconstruction method for 3D fluorescence tomography," *Optics Express*, vol. 13, no. 24, pp. 9847–9857, 2005.

[41] A. T. Kumar, S. B. Raymond, B. J. Bacskai, and D. A. Boas, "Comparison of frequency-domain and time-domain fluorescence lifetime tomography," *Optics letters*, vol. 33, no. 5, pp. 470–472, 2008.

[42] A. T. Kumar, S. B. Raymond, A. K. Dunn, B. J. Bacskai, and D. A. Boas, "A time domain fluorescence tomography system for small animal imaging," *IEEE transactions on medical imaging*, vol. 27, no. 8, pp. 1152–1163, 2008.

[43] F. Leblond, H. Dehghani, D. Kepshire, and B. W. Pogue, "Early-photon fluorescence tomography: spatial resolution improvements and noise stability considerations," *JOSA A*, vol. 26, no. 6, pp. 1444–1457, 2009.

[44] L. G. Henyey and J. L. Greenstein, "Diffuse radiation in the galaxy," *The Astrophysical Journal*, vol. 93, pp. 70–83, 1941.

[45] H. H. Barrett and K. J. Myers, *Foundations of image science.* John Wiley & Sons, 2013.

[46] S. R. Arridge, "Optical tomography in medical imaging," *Inverse problems*, vol. 15, no. 2, p. R41, 1999.

[47] S. Arridge, M. Schweiger, M. Hiraoka, and D. Delpy, "A finite element approach for modeling photon transport in tissue," *Medical physics*, vol. 20, no. 2, pp. 299–309, 1993.

[48] M. Schweiger and S. R. Arridge, "The toast++ software suite for forward and inverse modeling in optical tomography," *Journal of biomedical optics*, vol. 19, no. 4, p. 040801, 2014.

[49] S. R. Arridge and J. C. Schotland, "Optical tomography: forward and inverse problems," *Inverse problems*, vol. 25, no. 12, p. 123010, 2009.

[50] A. Gibson and H. Dehghani, "Diffuse optical imaging," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 367, no. 1900, pp. 3055–3072, 2009.

[51] A. Gibson, J. Hebden, and S. R. Arridge, "Recent advances in diffuse optical imaging," *Physics in Medicine & Biology*, vol. 50, no. 4, p. R1, 2005.

[52] A. D. Klose and E. W. Larsen, "Light transport in biological tissue based on the simplified spherical harmonics equations," *Journal of Computational Physics*, vol. 220, no. 1, pp. 441–470, 2006.

[53] A. H. Hielscher, R. E. Alcouffe, and R. L. Barbour, "Comparison of finite-difference transport and diffusion calculations for photon migration in homogeneous and heterogeneous tissues," *Physics in Medicine & Biology*, vol. 43, no. 5, p. 1285, 1998.

[54] A. K. Jha, M. A. Kupinski, T. Masumura, E. Clarkson, A. V. Maslov, and H. H. Barrett, "Simulating photon-transport in uniform media using the radiative

transport equation: a study using the Neumann-series approach," *JOSA A*, vol. 29, no. 8, pp. 1741–1757, 2012.

[55] A. K. Jha, M. A. Kupinski, H. H. Barrett, E. Clarkson, and J. H. Hartman, "Three-dimensional Neumann-series approach to model light transport in nonuniform media," *JOSA A*, vol. 29, no. 9, pp. 1885–1899, 2012.

[56] D. Dickey, O. Barajas, K. Brown, J. Tulip, and R. B. Moore, "Radiance modelling using the P3 approximation," *Physics in Medicine & Biology*, vol. 43, no. 12, p. 3559, 1998.

[57] E. Aydin, C. De Oliveira, and A. Goddard, "A finite element-spherical harmonics radiation transport model for photon migration in turbid media," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 84, no. 3, pp. 247–260, 2004.

[58] S. Wright, M. Schweiger, and S. Arridge, "Reconstruction in optical tomography using the PN approximations," *Measurement Science and Technology*, vol. 18, no. 1, p. 79, 2006.

[59] P. S. Mohan, T. Tarvainen, M. Schweiger, A. Pulkkinen, and S. R. Arridge, "Variable order spherical harmonic expansion scheme for the radiative transport equation using finite elements," *Journal of Computational Physics*, vol. 230, no. 19, pp. 7364–7383, 2011.

[60] Z. Yuan, X.-H. Hu, and H. Jiang, "A higher order diffusion model for three-dimensional photon migration and image reconstruction in optical tomography," *Physics in Medicine & Biology*, vol. 54, no. 1, p. 65, 2008.

[61] A. D. Klose, V. Ntziachristos, and A. H. Hielscher, "The inverse source problem based on the radiative transfer equation in optical molecular imaging," *Journal of Computational Physics*, vol. 202, no. 1, pp. 323–345, 2005.

[62] A. Pulkkinen and T. Tarvainen, "Truncated Fourier-series approximation of

the time-domain radiative transfer equation using finite elements," *J. Opt. Soc. Am. A*, vol. 30, no. 3, pp. 470–478, Mar 2013. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-30-3-470

[63] R. L. Harrison, "Introduction to Monte Carlo simulation," in *AIP conference proceedings*, vol. 1204, no. 1. American Institute of Physics, 2010, pp. 17–21.

[64] L. V. Wang and H.-i. Wu, *Biomedical optics: principles and imaging.* John Wiley & Sons, 2012.

[65] C. Zhu and Q. Liu, "Review of Monte Carlo modeling of light transport in tissues," *Journal of biomedical optics*, vol. 18, no. 5, p. 050902, 2013.

[66] J. Chen and X. Intes, "Time-gated perturbation Monte Carlo for whole body functional imaging in small animals," *Optics express*, vol. 17, no. 22, pp. 19 566–19 579, 2009.

[67] Y. Yuan, L. Yu, Z. Do*u*gan, and Q. Fang, "Graphics processing units-accelerated adaptive nonlocal means filter for denoising three-dimensional Monte Carlo photon transport simulations," *Journal of biomedical optics*, vol. 23, no. 12, p. 121618, 2018.

[68] Q. Fang and D. A. Boas, "Monte Carlo simulation of photon migration in 3D turbid media accelerated by graphics processing units," *Optics express*, vol. 17, no. 22, pp. 20 178–20 190, 2009.

[69] A. Jin, B. Yazıcı, and V. Ntziachristos, "Light illumination and detection patterns for fluorescence diffuse optical tomography based on compressive sensing," *IEEE transactions on image processing*, vol. 23, no. 6, pp. 2609–2624, 2014.

[70] S. C. Davis, H. Dehghani, J. Wang, S. Jiang, B. W. Pogue, and K. D. Paulsen, "Image-guided diffuse optical fluorescence tomography implemented with Laplacian-type regularization," *Optics express*, vol. 15, no. 7, pp. 4066–4082, 2007.

[71] X. Cao, B. Zhang, X. Wang, F. Liu, K. Liu, J. Luo, and J. Bai, “An adaptive Tikhonov regularization method for fluorescence molecular tomography,” *Medical & biological engineering & computing*, vol. 51, no. 8, pp. 849–858, 2013.

[72] L. Zhao, H. Yang, W. Cong, G. Wang, and X. Intes, “L_p regularization for early gate fluorescence molecular tomography,” *Optics letters*, vol. 39, no. 14, pp. 4156–4159, 2014.

[73] A. N. Tikhonov and V. Y. Arsenin, “Solutions of ill-posed problems,” *New York*, pp. 1–30, 1977.

[74] V. Ntziachristos, “Fluorescence molecular imaging,” *Annu. Rev. Biomed. Eng.*, vol. 8, pp. 1–33, 2006.

[75] J. Nocedal and S. Wright, *Numerical optimization.* Springer Science & Business Media, 2006.

[76] L. Zhang, F. Gao, H. He, and H. Zhao, “Three-dimensional scheme for time-domain fluorescence molecular tomography based on Laplace transforms with noise-robust factors,” *Optics express*, vol. 16, no. 10, pp. 7214–7223, 2008.

[77] P. C. Hansen, “The truncated SVD as a method for regularization,” *BIT Numerical Mathematics*, vol. 27, no. 4, pp. 534–553, 1987.

[78] H. E. Fleming, “Equivalence of regularization and truncated iteration in the solution of ill-posed image reconstruction problems,” *Linear Algebra and its applications*, vol. 130, pp. 133–150, 1990.

[79] H. Guo, J. Yu, X. He, Y. Hou, F. Dong, and S. Zhang, “Improved sparse reconstruction for fluorescence molecular tomography with $L_{1/2}$ regularization,” *Biomedical optics express*, vol. 6, no. 5, pp. 1648–1664, 2015.

[80] D. Zhu and C. Li, “Accelerated image reconstruction in fluorescence molecular tomography using a nonuniform updating scheme with momentum and ordered subsets methods,” *Journal of biomedical optics*, vol. 21, no. 1, p. 016004, 2016.

[81] ——, "Nonconvex regularizations in fluorescence molecular tomography for sparsity enhancement," *Physics in Medicine & Biology*, vol. 59, no. 12, p. 2901, 2014.

[82] A. Behrooz, H.-M. Zhou, A. A. Eftekhar, and A. Adibi, "Total variation regularization for 3D reconstruction in fluorescence tomography: experimental phantom studies," *Applied Optics*, vol. 51, no. 34, pp. 8216–8227, 2012.

[83] J. Dutta, S. Ahn, C. Li, S. R. Cherry, and R. M. Leahy, "Joint L_1 and total variation regularization for fluorescence molecular tomography," *Physics in Medicine & Biology*, vol. 57, no. 6, p. 1459, 2012.

[84] S. Jiang, J. Liu, Y. An, G. Zhang, J. Ye, Y. Mao, K. He, C. Chi, and J. Tian, "Novel $l_{2,1}$ norm optimization method for fluorescence molecular tomography reconstruction," *Biomedical optics express*, vol. 7, no. 6, pp. 2342–2359, 2016.

[85] S. Jiang, J. Liu, G. Zhang, Y. An, H. Meng, Y. Gao, K. Wang, and J. Tian, "Reconstruction of fluorescence molecular tomography via a fused LASSO method based on group sparsity prior," *IEEE Transactions on Biomedical Engineering*, vol. 66, no. 5, pp. 1361–1371, 2018.

[86] T. Correia, J. Aguirre, A. Sisniega, J. Chamorro-Servent, J. Abascal, J. J. Vaquero, M. Desco, V. Kolehmainen, and S. Arridge, "Split operator method for fluorescence diffuse optical tomography using anisotropic diffusion regularisation with prior anatomical information," *Biomedical optics express*, vol. 2, no. 9, pp. 2632–2648, 2011.

[87] J.-C. Baritaux, K. Hassler, M. Bucher, S. Sanyal, and M. Unser, "Sparsity-driven reconstruction for FDOT with anatomical priors," *IEEE Transactions on medical imaging*, vol. 30, no. 5, pp. 1143–1153, 2011.

[88] R. Baikejiang, Y. Zhao, B. Z. Fite, K. W. Ferrara, and C. Li, "Anatomical image-guided fluorescence molecular tomography reconstruction using kernel

method," *Journal of biomedical optics*, vol. 22, no. 5, p. 055001, 2017.

[89] D. Álvarez, P. Medina, and M. Moscoso, "Fluorescence lifetime imaging from time resolved measurements using a shape-based approach," *Optics Express*, vol. 17, no. 11, pp. 8843–8855, 2009.

[90] Y. An, J. Liu, G. Zhang, J. Ye, Y. Du, Y. Mao, C. Chi, and J. Tian, "A novel region reconstruction method for fluorescence molecular tomography," *IEEE Transactions on biomedical engineering*, vol. 62, no. 7, pp. 1818–1826, 2015.

[91] L. Guo, F. Liu, C. Cai, J. Liu, and G. Zhang, "3D deep encoder–decoder network for fluorescence molecular tomography," *Optics letters*, vol. 44, no. 8, pp. 1892–1895, 2019.

[92] H. Wang, J. Gao, Z. Zhao, X. Feng, W. Ma, J. Wang, and P. O. Ogunbona, "Fluorescence molecular tomography reconstruction of small targets using stacked auto-encoder neural networks," *IEEE Access*, vol. 8, pp. 37 657–37 663, 2020.

[93] R. Chandra and A. Rahmim, *Nuclear medicine physics: the basics.* Lippincott Williams & Wilkins, 2017.

[94] S. R. Cherry, J. A. Sorenson, and M. E. Phelps, *Physics in nuclear medicine.* Elsevier Health Sciences, 2012.

[95] A. Rahmim and H. Zaidi, "PET versus SPECT: strengths, limitations and challenges," *Nuclear medicine communications*, vol. 29, no. 3, pp. 193–207, 2008.

[96] S. R. Cherry, "In vivo molecular and genomic imaging: new challenges for imaging physics," *Physics in Medicine & Biology*, vol. 49, no. 3, p. R13, 2004.

[97] M. Soret, S. L. Bacharach, and I. Buvat, "Partial-volume effect in PET tumor imaging," *Journal of Nuclear Medicine*, vol. 48, no. 6, pp. 932–945, 2007.

[98] T. Kato, Y. Inui, A. Nakamura, and K. Ito, "Brain fluorodeoxyglucose (FDG) PET in dementia," *Ageing research reviews*, vol. 30, pp. 73–84, 2016.

[99] M. Schwaiger, S. Ziegler, and S. G. Nekolla, “PET/CT: challenge for nuclear cardiology,” *Journal of Nuclear Medicine*, vol. 46, no. 10, pp. 1664–1678, 2005.

[100] L. Jødal, C. Le Loirec, and C. Champion, “Positron range in PET imaging: an alternative approach for assessing and correcting the blurring,” *Physics in Medicine & Biology*, vol. 57, no. 12, p. 3931, 2012.

[101] T. K. Lewellen, “Recent developments in PET detector technology,” *Physics in Medicine & Biology*, vol. 53, no. 17, p. R287, 2008.

[102] R. Nutt, “The history of positron emission tomography,” *Molecular Imaging & Biology*, vol. 4, no. 1, pp. 11–26, 2002.

[103] M. E. Phelps and S. R. Cherry, “The changing design of positron imaging systems,” *Clinical Positron Imaging*, vol. 1, no. 1, pp. 31–45, 1998.

[104] https://depts.washington.edu/imreslab/from%20old%20SITE/pet_intro/intro_src/section5.html.

[105] P. Zanzonico, “Positron emission tomography: a review of basic principles, scanner design and performance, and current systems,” in *Seminars in nuclear medicine*, vol. 34, no. 2. WB Saunders, 2004, pp. 87–111.

[106] W. W. Moses, “Recent advances and future advances in time-of-flight PET,” in *AIP Conference Proceedings*, vol. 1204, no. 1. American Institute of Physics, 2010, pp. 119–125.

[107] M. Miller, J. Zhang, K. Binzel, J. Griesmer, T. Laurence, M. Narayanan, D. Natarajamani, S. Wang, and M. Knopp, “Characterization of the vereos digital photon counting PET system,” *Journal of Nuclear Medicine*, vol. 56, no. supplement 3, pp. 434–434, 2015.

[108] J. van Sluis, J. de Jong, J. Schaar, W. Noordzij, P. van Snick, R. Dierckx, R. Borra, A. Willemsen, and R. Boellaard, “Performance characteristics of the digital Biograph Vision PET/CT system,” *Journal of Nuclear Medicine*, vol. 60,

no. 7, pp. 1031–1036, 2019.

[109] V. Bettinardi, L. Presotto, E. Rapisarda, M. Picchio, L. Gianolli, and M. Gilardi, “Physical performance of the new hybrid PET/CT Discovery-690,” *Medical physics*, vol. 38, no. 10, pp. 5394–5411, 2011.

[110] M. Conti, “Focus on time-of-flight PET: the benefits of improved time resolution,” *European journal of nuclear medicine and molecular imaging*, vol. 38, no. 6, pp. 1147–1157, 2011.

[111] S. Vandenberghe, E. Mikhaylova, E. D’Hoe, P. Mollet, and J. S. Karp, “Recent developments in time-of-flight PET,” *EJNMMI physics*, vol. 3, no. 1, p. 3, 2016.

[112] S. Tong, A. M. Alessio, and P. E. Kinahan, “Image reconstruction for PET/CT scanners: past achievements and future challenges,” *Imaging in medicine*, vol. 2, no. 5, p. 529, 2010.

[113] M. A. Lodge, R. D. Badawi, R. Gilbert, P. E. Dibos, and B. R. Line, “Comparison of 2-dimensional and 3-dimensional acquisition for 18F-FDG PET oncology studies performed on an LSO-based scanner,” *Journal of Nuclear Medicine*, vol. 47, no. 1, pp. 23–31, 2006.

[114] S. R. Cherry, M. Dahlbom, and E. J. Hoffman, “3D PET using a conventional multislice tomograph without septa.” *Journal of computer assisted tomography*, vol. 15, no. 4, pp. 655–668, 1991.

[115] D. L. Bailey, M. P. Miller, T. J. Spinks, P. M. Bloomfield, L. Livieratos, H. E. Young, and T. Jones, “Experience with fully 3D PET and implications for future high-resolution 3D tomographs,” *Physics in Medicine & Biology*, vol. 43, no. 4, p. 777, 1998.

[116] P. Markiewicz, K. Thielemans, J. Schott, D. Atkinson, S. Arridge, B. Hutton, and S. Ourselin, “Rapid processing of PET list-mode data for efficient uncertainty estimation and data analysis,” *Physics in Medicine & Biology*, vol. 61, no. 13, p.

N322, 2016.

[117] A. Rahmim, M. Lenox, A. J. Reader, C. Michel, Z. Burbar, T. J. Ruth, and V. Sossi, "Statistical list-mode image reconstruction for the high resolution research tomograph," *Physics in Medicine & Biology*, vol. 49, no. 18, p. 4239, 2004.

[118] F. H. Fahey, "Data acquisition in PET imaging," *Journal of nuclear medicine technology*, vol. 30, no. 2, pp. 39–49, 2002.

[119] V. Kapoor, B. M. McCook, and F. S. Torok, "An introduction to PET-CT imaging," *Radiographics*, vol. 24, no. 2, pp. 523–543, 2004.

[120] P. E. Kinahan, D. Townsend, T. Beyer, and D. Sashin, "Attenuation correction for a combined 3D PET/CT scanner," *Medical physics*, vol. 25, no. 10, pp. 2046–2053, 1998.

[121] I. Bezrukov, F. Mantlik, H. Schmidt, B. Schölkopf, and B. J. Pichler, "MR-based PET attenuation correction for PET/MR imaging," in *Seminars in nuclear medicine*, vol. 43, no. 1. Elsevier, 2013, pp. 45–59.

[122] O. Rousset, A. Rahmim, A. Alavi, and H. Zaidi, "Partial volume correction strategies in PET," *PET clinics*, vol. 2, no. 2, pp. 235–249, 2007.

[123] J. L. Prince and J. M. Links, *Medical imaging signals and systems.* Pearson Prentice Hall Upper Saddle River, NJ, 2006.

[124] J. Hsieh, *Computed tomography: principles, design, artifacts, and recent advances.* SPIE press, 2003, vol. 114.

[125] H. M. Hudson and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE transactions on medical imaging*, vol. 13, no. 4, pp. 601–609, 1994.

[126] S. Tong, A. Alessio, and P. Kinahan, "Noise and signal properties in PSF-based

fully 3D PET image reconstruction: an experimental evaluation," *Physics in Medicine & Biology*, vol. 55, no. 5, p. 1453, 2010.

[127] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 721–741, 1984.

[128] P. J. Green, "On use of the EM algorithm for penalized likelihood estimation," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 52, no. 3, pp. 443–452, 1990.

[129] ——, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE transactions on medical imaging*, vol. 9, no. 1, pp. 84–93, 1990.

[130] J. A. Fessler and A. O. Hero, "Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms," *IEEE Transactions on Image Processing*, vol. 4, no. 10, pp. 1417–1429, 1995.

[131] ——, "Space-alternating generalized expectation-maximization algorithm," *IEEE Transactions on signal processing*, vol. 42, no. 10, pp. 2664–2677, 1994.

[132] A. R. De Pierro and M. B. Yamagishi, "Fast EM-like methods for maximum" a posteriori" estimates in emission tomography," *IEEE transactions on medical imaging*, vol. 20, no. 4, pp. 280–288, 2001.

[133] H. Erdogan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Physics in Medicine & Biology*, vol. 44, no. 11, p. 2835, 1999.

[134] E. U. Mumcuoglu, R. Leahy, S. R. Cherry, and Z. Zhou, "Fast gradient-based methods for Bayesian reconstruction of transmission and emission PET images," *IEEE transactions on Medical Imaging*, vol. 13, no. 4, pp. 687–701, 1994.

[135] Y.-J. Tsai, A. Bousse, M. J. Ehrhardt, C. W. Stearns, S. Ahn, B. F. Hutton, S. Arridge, and K. Thielemans, "Fast quasi-newton algorithms for penalized

reconstruction in emission tomography and further improvements via preconditioning," *IEEE transactions on medical imaging*, vol. 37, no. 4, pp. 1000–1010, 2017.

[136] D. L. Bailey, "Transmission scanning in emission tomography," *European journal of nuclear medicine*, vol. 25, no. 7, pp. 774–787, 1998.

[137] S. R. Cherry and S.-C. Huang, "Effects of scatter on model parameter estimates in 3D PET studies of the human brain," *IEEE transactions on nuclear science*, vol. 42, no. 4, pp. 1174–1179, 1995.

[138] S. Grootoonk, T. Spinks, D. Sashin, N. Spyrou, and T. Jones, "Correction for scatter in 3D brain PET using a dual energy window method," *Physics in Medicine & Biology*, vol. 41, no. 12, p. 2757, 1996.

[139] L. Shao, R. Freifelder, and J. S. Karp, "Triple energy window scatter correction technique in PET," *IEEE transactions on medical imaging*, vol. 13, no. 4, pp. 641–648, 1994.

[140] J. M. Ollinger, "Model-based scatter correction for fully 3D PET," *Physics in Medicine & Biology*, vol. 41, no. 1, p. 153, 1996.

[141] C. C. Watson, D. Newport, and M. E. Casey, "A single scatter simulation technique for scatter correction in 3D PET," in *Three-dimensional image reconstruction in radiology and nuclear medicine.* Springer, 1996, pp. 255–268.

[142] R. D. Badawi and P. Marsden, "Developments in component-based normalization for 3D PET," *Physics in Medicine & Biology*, vol. 44, no. 2, p. 571, 1999.

[143] L. Theodorakis, G. Loudos, V. Prassopoulos, C. Kappas, I. Tsougos, and P. Georgoulias, "A review of PET normalization: striving for count rate uniformity," *Nuclear medicine communications*, vol. 34, no. 11, pp. 1033–1045, 2013.

[144] G. Germano and E. J. Hoffman, "Investigation of count rate and deadtime characteristics of a high resolution PET system." *Journal of computer assisted*

tomography, vol. 12, no. 5, pp. 836–846, 1988.

[145] D. L. Bailey, M. N. Maisey, D. W. Townsend, and P. E. Valk, *Positron emission tomography.* Springer, 2005, vol. 2.

[146] M. V. Green, H. G. Ostrow, J. Seidel, and M. G. Pomper, "Experimental evaluation of depth-of-interaction correction in a small-animal positron emission tomography scanner," *Molecular imaging*, vol. 9, no. 6, pp. 7290–2010, 2010.

[147] C. S. Levin, "Design of a high-resolution and high-sensitivity scintillation crystal array for PET with nearly complete light collection," *IEEE Transactions on Nuclear Science*, vol. 49, no. 5, pp. 2236–2243, 2002.

[148] R. Wiener, S. Surti, and J. Karp, "DOI determination by rise time discrimination in single-ended readout for TOF PET imaging," *IEEE transactions on nuclear science*, vol. 60, no. 3, pp. 1478–1486, 2013.

[149] J. P. Schmall, J. S. Karp, M. Werner, and S. Surti, "Parallax error in long-axial field-of-view PET scanners—a simulation study," *Physics in Medicine & Biology*, vol. 61, no. 14, p. 5443, 2016.

[150] M. Pizzichemi, G. Stringhini, T. Niknejad, Z. Liu, P. Lecoq, S. Tavernier, J. Varela, M. Paganoni, and E. Auffray, "A new method for depth of interaction determination in PET detectors," *Physics in Medicine & Biology*, vol. 61, no. 12, p. 4679, 2016.

[151] A. Rahmim, J. Qi, and V. Sossi, "Resolution modeling in PET imaging: theory, practice, benefits, and pitfalls," *Medical physics*, vol. 40, no. 6Part1, 2013.

[152] Y. Lu, J.-D. Gallezot, M. Naganawa, S. Ren, K. Fontaine, J. Wu, J. A. Onofrey, T. Toyonaga, N. Boutagy, T. Mulnix *et al.*, "Data-driven voluntary body motion detection and non-rigid event-by-event correction for static and dynamic PET," *Physics in Medicine & Biology*, vol. 64, no. 6, p. 065002, 2019.

[153] A. Rahmim, O. Rousset, and H. Zaidi, "Strategies for motion tracking and

correction in PET," *PET clinics*, vol. 2, no. 2, pp. 251–266, 2007.

[154] W. W. Moses, "Fundamental limits of spatial resolution in PET," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 648, pp. S236–S240, 2011.

[155] B. F. Hutton and A. Osiecki, "Correction of partial volume effects in myocardial SPECT," *Journal of Nuclear Cardiology*, vol. 5, no. 4, pp. 402–413, 1998.

[156] K. Erlandsson, I. Buvat, P. H. Pretorius, B. A. Thomas, and B. F. Hutton, "A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology," *Physics in Medicine & Biology*, vol. 57, no. 21, p. R119, 2012.

[157] M. A. G. Ballester, A. P. Zisserman, and M. Brady, "Estimation of the partial volume effect in MRI," *Medical image analysis*, vol. 6, no. 4, pp. 389–405, 2002.

[158] J. Dehmeshki, X. Ye, H. Amin, M. Abaei, X. Lin, and S. D. Qanadli, "Volumetric quantification of atherosclerotic plaque in CT considering partial volume effect," *IEEE Transactions on Medical Imaging*, vol. 26, no. 3, pp. 273–282, 2007.

[159] P. Pretorius, M. King, T. Pan, D. De Vries, S. Glick, and C. Byrne, "Reducing the influence of the partial volume effect on SPECT activity quantitation with 3D modelling of spatial resolution in iterative reconstruction," *Physics in Medicine & Biology*, vol. 43, no. 2, p. 407, 1998.

[160] H. Mohy-ud Din, A. Nicolas, W. Willis, K. Abdel, F. Dean, A. Rahmim *et al.*, "Intra-frame motion compensation in multi-frame brain PET imaging," *Frontiers in Biomedical Technologies*, vol. 2, no. 2, pp. 60–72, 2015.

[161] A. Rahmim, K. Dinelle, J.-C. Cheng, M. A. Shilov, W. P. Segars, S. C. Lidstone, S. Blinder, O. G. Rousset, H. Vajihollahi, B. M. Tsui *et al.*, "Accurate event-driven motion compensation in high-resolution PET incorporating scattered and random events," *IEEE Transactions on Medical Imaging*, vol. 27, no. 8, pp.

1018–1033, 2008.

[162] A. Rahmim, P. Bloomfield, S. Houle, M. Lenox, C. Michel, K. R. Buckley, T. J. Ruth, and V. Sossi, "Motion compensation in histogram-mode and list-mode EM reconstructions: beyond the event-driven approach," *IEEE Transactions on Nuclear Science*, vol. 51, no. 5, pp. 2588–2596, 2004.

[163] J. Ouyang, Q. Li, and G. El Fakhri, "Magnetic resonance-based motion correction for positron emission tomography imaging," in *Seminars in nuclear medicine*, vol. 43, no. 1. Elsevier, 2013, pp. 60–67.

[164] M. G. Ullisch, J. J. Scheins, C. Weirich, E. R. Kops, A. Celik, L. Tellmann, T. Stöcker, H. Herzog, and N. J. Shah, "Mr-based PET motion correction procedure for simultaneous MR-PET neuroimaging of human brain," *PloS one*, vol. 7, no. 11, 2012.

[165] C. Tsoumpas, J. E. Mackewn, P. Halsted, A. P. King, C. Buerger, J. J. Totman, T. Schaeffter, and P. K. Marsden, "Simultaneous PET–MR acquisition and MR-derived motion fields for correction of non-rigid motion in pet," *Annals of nuclear medicine*, vol. 24, no. 10, pp. 745–750, 2010.

[166] Y. Picard and C. J. Thompson, "Motion correction of PET images using multiple acquisition frames," *IEEE transactions on medical imaging*, vol. 16, no. 2, pp. 137–144, 1997.

[167] H. Zaidi, *Quantitative analysis in nuclear medicine imaging.* Springer, 2006.

[168] D. Gutierrez, M.-L. Montandon, F. Assal, M. Allaoua, O. Ratib, K.-O. Lövblad, and H. Zaidi, "Anatomically guided voxel-based partial volume effect correction in brain PET: impact of MRI segmentation," *Computerized Medical Imaging and Graphics*, vol. 36, no. 8, pp. 610–619, 2012.

[169] D. L. Barbee, R. T. Flynn, J. E. Holden, R. J. Nickles, and R. Jeraj, "A method for partial volume correction of PET-imaged tumor heterogeneity using expecta-

tion maximization with a spatially varying point spread function," *Physics in Medicine & Biology*, vol. 55, no. 1, p. 221, 2009.

[170] D. A. Boas, A. M. Dale, and M. A. Franceschini, "Diffuse optical imaging of brain activation: approaches to optimizing image sensitivity, resolution, and accuracy," *Neuroimage*, vol. 23, pp. S275–S288, 2004.

[171] G. Yu, T. Durduran, G. Lech, C. Zhou, B. Chance, E. R. Mohler, and A. G. Yodh, "Time-dependent blood flow and oxygenation in human skeletal muscles measured with noninvasive near-infrared diffuse optical spectroscopies," *Journal of biomedical optics*, vol. 10, no. 2, p. 024027, 2005.

[172] A. Doronin and I. Meglinski, "Peer-to-peer Monte Carlo simulation of photon migration in topical applications of biomedical optics," *Journal of Biomedical Optics*, vol. 17, no. 9, p. 090504, 2012.

[173] W. M. Star, "Light dosimetry in vivo," *Physics in Medicine & Biology*, vol. 42, no. 5, p. 763, 1997.

[174] E. Aydin, "Three-dimensional photon migration through voidlike regions and channels," *Applied optics*, vol. 46, no. 34, pp. 8272–8277, 2007.

[175] A. H. Hielscher and R. E. Alcouffe, "Discrete-ordinate transport simulations of light propagation in highly forward scattering heterogeneous media," in *Advances in Optical Imaging and Photon Migration.* Optical Society of America, 1998, p. ATuC2.

[176] G. W. Faris, "Diffusion equation boundary conditions for the interface between turbid media: a comment," *JOSA A*, vol. 19, no. 3, pp. 519–520, 2002.

[177] R. C. Haskell, L. O. Svaasand, T.-T. Tsay, T.-C. Feng, M. S. McAdams, and B. J. Tromberg, "Boundary conditions for the diffusion equation in radiative transfer," *JOSA A*, vol. 11, no. 10, pp. 2727–2741, 1994.

[178] M. Machida, G. Y. Panasyuk, J. C. Schotland, and V. A. Markel, "The Green's

function for the radiative transport equation in the slab geometry," *Journal of Physics A: Mathematical and Theoretical*, vol. 43, no. 6, p. 065402, 2010.

[179] A. Liemert and A. Kienle, "Analytical approach for solving the radiative transfer equation in two-dimensional layered media," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 113, no. 7, pp. 559–564, 2012.

[180] ——, "Exact and efficient solution of the radiative transport equation for the semi-infinite medium," *Scientific reports*, vol. 3, no. 1, pp. 1–7, 2013.

[181] R. Aronson, "Boundary conditions for diffusion of light," *JOSA A*, vol. 12, no. 11, pp. 2532–2539, 1995.

[182] A. Kienle and M. S. Patterson, "Improved solutions of the steady-state and the time-resolved diffusion equations for reflectance from a semi-infinite turbid medium," *JOSA A*, vol. 14, no. 1, pp. 246–254, 1997.

[183] A. D. Kim, "Correcting the diffusion approximation at the boundary," *JOSA A*, vol. 28, no. 6, pp. 1007–1015, 2011.

[184] O. Lehtikangas, T. Tarvainen, and A. D. Kim, "Modeling boundary measurements of scattered light using the corrected diffusion approximation," *Biomedical optics express*, vol. 3, no. 3, pp. 552–571, 2012.

[185] M. Schweiger, S. Arridge, M. Hiraoka, and D. Delpy, "The finite element method for the propagation of light in scattering media: boundary and source conditions," *Medical physics*, vol. 22, no. 11, pp. 1779–1792, 1995.

[186] A. K. Jha, E. Clarkson, and M. A. Kupinski, "An ideal-observer framework to investigate signal detectability in diffuse optical imaging," *Biomedical optics express*, vol. 4, no. 10, pp. 2107–2123, 2013.

[187] T. Tarvainen, M. Vauhkonen, V. Kolehmainen, S. R. Arridge, and J. P. Kaipio, "Coupled radiative transfer equation and diffusion approximation model for photon migration in turbid medium with low-scattering and non-scattering

regions," *Physics in Medicine & Biology*, vol. 50, no. 20, p. 4913, 2005.

[188] L. Wang and S. L. Jacques, "Hybrid model of Monte Carlo simulation and diffusion theory for light reflectance by turbid media," *JOSA A*, vol. 10, no. 8, pp. 1746–1752, 1993.

[189] O. Lehtikangas, T. Tarvainen, A. Kim, and S. R. Arridge, "Finite element approximation of the radiative transport equation in a medium with piece-wise constant refractive index," *Journal of Computational Physics*, vol. 282, pp. 345–359, 2015.

[190] C. K. Hayakawa, J. Spanier, F. Bevilacqua, A. K. Dunn, J. S. You, B. J. Tromberg, and V. Venugopalan, "Perturbation Monte Carlo methods to solve inverse photon migration problems in heterogeneous tissues," *Optics letters*, vol. 26, no. 17, pp. 1335–1337, 2001.

[191] Y. P. Kumar and R. M. Vasu, "Reconstruction of optical properties of low-scattering tissue using derivative estimated through perturbation Monte-Carlo method," *Journal of biomedical optics*, vol. 9, no. 5, pp. 1002–1013, 2004.

[192] R. Yao, X. Intes, and Q. Fang, "Direct approach to compute Jacobians for diffuse optical tomography using perturbation Monte Carlo-based photon "replay"," *Biomedical optics express*, vol. 9, no. 10, pp. 4588–4603, 2018.

[193] N. C. Biswal, A. Aguirre, Y. Xu, S. Zanganeh, Q. Zhu, C. Pavlik, M. B. Smith, L. T. Kuhn, and K. P. Claffey, "Imaging tumor hypoxia by near-infrared fluorescence tomography," *Journal of biomedical optics*, vol. 16, no. 6, p. 066009, 2011.

[194] L. Zhou and B. Yazici, "Discretization error analysis and adaptive meshing algorithms for fluorescence diffuse optical tomography in the presence of measurement noise," *IEEE transactions on image processing*, vol. 20, no. 4, pp. 1094–1111, 2010.

[195] V. Ntziachristos and R. Weissleder, “Experimental three-dimensional fluorescence reconstruction of diffuse media by use of a normalized born approximation,” *Optics letters*, vol. 26, no. 12, pp. 893–895, 2001.

[196] J. Ye, C. Chi, Z. Xue, P. Wu, Y. An, H. Xu, S. Zhang, and J. Tian, “Fast and robust reconstruction for fluorescence molecular tomography via a sparsity adaptive subspace pursuit method,” *Biomedical optics express*, vol. 5, no. 2, pp. 387–406, 2014.

[197] D. Han, J. Tian, S. Zhu, J. Feng, C. Qin, B. Zhang, and X. Yang, “A fast reconstruction algorithm for fluorescence molecular tomography with sparsity regularization,” *Optics express*, vol. 18, no. 8, pp. 8630–8646, 2010.

[198] O. Lee, J. M. Kim, Y. Bresler, and J. C. Ye, “Compressive diffuse optical tomography: noniterative exact reconstruction using joint sparsity,” *IEEE transactions on medical imaging*, vol. 30, no. 5, pp. 1129–1142, 2011.

[199] A. M. Bruckstein, D. L. Donoho, and M. Elad, “From sparse solutions of systems of equations to sparse modeling of signals and images,” *SIAM review*, vol. 51, no. 1, pp. 34–81, 2009.

[200] V. Ntziachristos and R. Weissleder, “Charge-coupled-device based scanner for tomography of fluorescent near-infrared probes in turbid media,” *Medical Physics*, vol. 29, no. 5, pp. 803–809, 2002.

[201] L. Adhikari, D. Zhu, C. Li, and R. F. Marcia, “Nonconvex reconstruction for low-dimensional fluorescence molecular tomographic poisson observations,” in *2015 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2015, pp. 2404–2408.

[202] Q. Pian, R. Yao, L. Zhao, and X. Intes, “Hyperspectral time-resolved wide-field fluorescence molecular tomography based on structured light and single-pixel detection,” *Optics letters*, vol. 40, no. 3, pp. 431–434, 2015.

[203] Y. Vardi, L. Shepp, and L. Kaufman, “A statistical model for positron emission tomography,” *Journal of the American statistical Association*, vol. 80, no. 389, pp. 8–20, 1985.

[204] K. Lange, M. Bahn, and R. Little, “A theoretical study of some maximum likelihood algorithms for emission and transmission tomography,” *IEEE Transactions on Medical Imaging*, vol. 6, no. 2, pp. 106–114, 1987.

[205] J. Llacer, E. Veklerov, K. J. Coakley, E. J. Hoffman, and J. Nunez, “Statistical analysis of maximum likelihood estimator images of human brain FDG PET studies,” *IEEE transactions on medical imaging*, vol. 12, no. 2, pp. 215–231, 1993.

[206] R. M. Lewitt and S. Matej, “Overview of methods for image reconstruction from projections in emission computed tomography,” *Proceedings of the IEEE*, vol. 91, no. 10, pp. 1588–1611, 2003.

[207] G. Alexandrakis, F. R. Rannou, and A. F. Chatziioannou, “Tomographic bioluminescence imaging by use of a combined optical-PET (OPET) system: a computer simulation feasibility study,” *Physics in Medicine & Biology*, vol. 50, no. 17, p. 4225, 2005.

[208] N. V. Slavine, M. A. Lewis, E. Richer, and P. P. Antich, “Iterative reconstruction method for light emitting sources based on the diffusion equation,” *Medical physics*, vol. 33, no. 1, pp. 61–68, 2006.

[209] M. Jiang, T. Zhou, J. Cheng, W. Cong, and G. Wang, “Image reconstruction for bioluminescence tomography from partial measurement,” *Optics Express*, vol. 15, no. 18, pp. 11 095–11 116, 2007.

[210] L. Cao and J. Peter, “Bayesian reconstruction strategy of fluorescence-mediated tomography using an integrated SPECT-CT-OT system,” *Physics in Medicine & Biology*, vol. 55, no. 9, p. 2693, 2010.

[211] J. Qi and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Physics in Medicine & Biology*, vol. 51, no. 15, p. R541, 2006.

[212] S. Ahn, A. J. Chaudhari, F. Darvas, C. A. Bouman, and R. M. Leahy, "Fast iterative image reconstruction methods for fully 3D multispectral bioluminescence tomography," *Physics in Medicine & Biology*, vol. 53, no. 14, p. 3921, 2008.

[213] D. Ma, P. Wolf, A. V. Clough, and T. G. Schmidt, "The performance of MLEM for dynamic imaging from simulated few-view, multi-pinhole SPECT," *IEEE transactions on nuclear science*, vol. 60, no. 1, pp. 115–123, 2012.

[214] G. Kontaxakis and L. G. Strauss, "Maximum likelihood algorithms for image reconstruction in positron emission tomography," *Radionuclides for Oncology*, pp. 73–106, 1998.

[215] J.-H. Chang, J. M. Anderson, and J. R. Votaw, "Regularized image reconstruction algorithms for positron emission tomography," *IEEE transactions on medical imaging*, vol. 23, no. 9, pp. 1165–1175, 2004.

[216] A. Jin, B. Yazici, A. Ale, and V. Ntziachristos, "Preconditioning of the fluorescence diffuse optical tomography sensing matrix based on compressive sensing," *Optics letters*, vol. 37, no. 20, pp. 4326–4328, 2012.

[217] J. Shi, X. Cao, F. Liu, B. Zhang, J. Luo, and J. Bai, "Greedy reconstruction algorithm for fluorescence molecular tomography by means of truncated singular value decomposition conversion," *JOSA A*, vol. 30, no. 3, pp. 437–447, 2013.

[218] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM journal on imaging sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[219] G. Strangman, M. A. Franceschini, and D. A. Boas, "Factors affecting the accuracy of near-infrared spectroscopy concentration calculations for focal changes in oxygenation parameters," *Neuroimage*, vol. 18, no. 4, pp. 865–879, 2003.

[220] W. P. Segars, B. M. Tsui, E. C. Frey, G. A. Johnson, and S. S. Berr, “Development of a 4-D digital mouse phantom for molecular imaging research,” *Molecular Imaging & Biology*, vol. 6, no. 3, pp. 149–159, 2004.

[221] X. Liu, F. Liu, Y. Zhang, and J. Bai, “Unmixing dynamic fluorescence diffuse optical tomography images with independent component analysis,” *IEEE Transactions on medical imaging*, vol. 30, no. 9, pp. 1591–1604, 2011.

[222] A. K. Jha, “Retrieving information from scattered photons in medical imaging,” 2013.

[223] M. Rullmann, J. Dukart, K.-T. Hoffmann, J. Luthardt, S. Tiepolt, M. Patt, H.-J. Gertz, M. L. Schroeter, J. Seibyl, W. J. Schulz-Schaeffer *et al.*, “Partial-volume effect correction improves quantitative analysis of 18F-florbetaben β-amyloid PET scans,” *Journal of Nuclear Medicine*, vol. 57, no. 2, pp. 198–203, 2016.

[224] J. O. Rinne, R. Portin, H. Ruottinen, E. Nurmi, J. Bergman, M. Haaparanta, and O. Solin, “Cognitive impairment and the brain dopaminergic system in Parkinson disease:[18F] fluorodopa positron emission tomographic study,” *Archives of neurology*, vol. 57, no. 4, pp. 470–475, 2000.

[225] K. Erlandsson, J. Dickson, S. Arridge, D. Atkinson, S. Ourselin, and B. F. Hutton, “MR imaging–guided partial volume correction of PET data in PET/MR imaging,” *PET clinics*, vol. 11, no. 2, pp. 161–177, 2016.

[226] F. Kotasidis, J. Matthews, G. Angelis, P. Noonan, A. Jackson, P. Price, W. Lionheart, and A. Reader, “Single scan parameterization of space-variant point spread functions in image space via a printed array: the impact for two PET/CT scanners,” *Physics in Medicine & Biology*, vol. 56, no. 10, p. 2917, 2011.

[227] V. Y. Panin, F. Kehren, C. Michel, and M. Casey, “Fully 3-D PET reconstruction with system matrix derived from point source measurements,” *IEEE transactions on medical imaging*, vol. 25, no. 7, pp. 907–921, 2006.

[228] A. J. Reader, S. Ally, F. Bakatselos, R. Manavaki, R. J. Walledge, A. P. Jeavons, P. J. Julyan, S. Zhao, D. L. Hastings, and J. Zweit, “One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays,” *IEEE Transactions on Nuclear Science*, vol. 49, no. 3, pp. 693–699, 2002.

[229] J. Dutta, G. El Fakhri, X. Zhu, and Q. Li, “PET point spread function modeling and image deblurring using a PET/MRI joint entropy prior,” in *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2015, pp. 1423–1426.

[230] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar, “High-resolution 3D bayesian image reconstruction using the microPET small-animal scanner,” *Physics in medicine & biology*, vol. 43, no. 4, p. 1001, 1998.

[231] A. Iriarte, R. Marabini, S. Matej, C. O. S. Sorzano, and R. M. Lewitt, “System models for PET statistical iterative reconstruction: A review,” *Computerized Medical Imaging and Graphics*, vol. 48, pp. 30–48, 2016.

[232] S. Ashrafinia, H. Mohy-ud Din, N. A. Karakatsanis, A. K. Jha, M. E. Casey, D. J. Kadrmas, and A. Rahmim, “Generalized PSF modeling for optimized quantitation in PET imaging,” *Physics in Medicine & Biology*, vol. 62, no. 12, p. 5149, 2017.

[233] A. Mehranian, M. A. Belzunce, F. Niccolini, M. Politis, C. Prieto, F. Turkheimer, A. Hammers, and A. J. Reader, “PET image reconstruction using multi-parametric anato-functional priors,” *Physics in Medicine & Biology*, vol. 62, no. 15, p. 5975, 2017.

[234] O. G. Rousset, Y. Ma, A. C. Evans *et al.*, “Correction for partial volume effects in PET: principle and validation,” *Journal of Nuclear Medicine*, vol. 39, no. 5, pp. 904–911, 1998.

[235] C. LABBÉ, M. KOEPP, J. ASHBURNER, T. SPINKS, M. RICHARDSON,

J. DUNCAN, and V. CUNNINGHAM, “Absolute PET quantification with correction for partial volume effects within cerebral structures,” in *Quantitative functional brain imaging with positron emission tomography.* Elsevier, 1998, pp. 59–66.

[236] M. Sattarivand, M. Kusano, I. Poon, and C. Caldwell, “Symmetric geometric transfer matrix partial volume correction for PET imaging: principle, validation and robustness,” *Physics in Medicine & Biology*, vol. 57, no. 21, p. 7101, 2012.

[237] Y. Du, B. M. Tsui, and E. C. Frey, “Partial volume effect compensation for quantitative brain SPECT imaging,” *IEEE transactions on medical imaging*, vol. 24, no. 8, pp. 969–976, 2005.

[238] H. W. Müller-Gärtner, J. M. Links, J. L. Prince, R. N. Bryan, E. McVeigh, J. P. Leal, C. Davatzikos, and J. J. Frost, “Measurement of radiotracer concentration in brain gray matter using positron emission tomography: MRI-based correction for partial volume effects,” *Journal of Cerebral Blood Flow & Metabolism*, vol. 12, no. 4, pp. 571–583, 1992.

[239] C. C. Meltzer, J. K. Zubieta, J. M. Links, P. Brakeman, M. J. Stumpf, and J. J. Frost, “MR-based correction of brain PET measurements for heterogeneous gray matter radioactivity distribution,” *Journal of Cerebral Blood Flow & Metabolism*, vol. 16, no. 4, pp. 650–658, 1996.

[240] K. Erlandsson, A. Wong, R. Van Heertum, J. Mann, and R. Parsey, “An improved method for voxel-based partial volume correction in PET and SPECT,” *Neuroimage*, vol. 31, no. Supplement 2, p. T84, 2006.

[241] B. A. Thomas, V. Cuplov, A. Bousse, A. Mendes, K. Thielemans, B. F. Hutton, and K. Erlandsson, “PETPVC: a toolbox for performing partial volume correction techniques in positron emission tomography,” *Physics in Medicine & Biology*, vol. 61, no. 22, p. 7975, 2016.

[242] J. Yang, S. Huang, M. Mega, K. Lin, A. Toga, G. Small, and M. Phelps, “Investigation of partial volume correction methods for brain FDG PET studies,” *IEEE Transactions on Nuclear Science*, vol. 43, no. 6, pp. 3322–3327, 1996.

[243] B. A. Thomas, K. Erlandsson, M. Modat, L. Thurfjell, R. Vandenberghe, S. Ourselin, and B. F. Hutton, “The importance of appropriate partial volume correction for PET quantification in Alzheimer’s disease,” *European journal of nuclear medicine and molecular imaging*, vol. 38, no. 6, pp. 1104–1119, 2011.

[244] J. Tohka and A. Reilhac, “Deconvolution-based partial volume correction in Raclopride-PET and Monte Carlo comparison to MR-based method,” *Neuroimage*, vol. 39, no. 4, pp. 1570–1584, 2008.

[245] S. S. Golla, M. Lubberink, B. N. van Berckel, A. A. Lammertsma, and R. Boellaard, “Partial volume correction of brain PET studies using iterative deconvolution in combination with HYPR denoising,” *EJNMMI research*, vol. 7, no. 1, p. 36, 2017.

[246] A. S. Kirov, J. Piao, and C. R. Schmidtlein, “Partial volume effect correction in PET using regularized iterative deconvolution with variance control based on local topology,” *Physics in Medicine & Biology*, vol. 53, no. 10, p. 2577, 2008.

[247] N. Boussion, C. C. Le Rest, M. Hatt, and D. Visvikis, “Incorporation of wavelet-based denoising in iterative deconvolution for partial volume correction in whole-body PET imaging,” *European journal of nuclear medicine and molecular imaging*, vol. 36, no. 7, pp. 1064–1075, 2009.

[248] Y. Gao, H. Zhang, Y. Zhu, M. Bilgel, O. Rousset, S. Resnick, D. F. Wong, L. Lu, and A. Rahmim, “Voxel-based partial volume correction of amyloid PET images incorporating non-local means regularization,” in *2018 IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings (NSS/MIC)*. IEEE, 2018, pp. 1–4.

[249] M. J. Ehrhardt, K. Thielemans, L. Pizarro, D. Atkinson, S. Ourselin, B. F. Hutton, and S. R. Arridge, "Joint reconstruction of PET-MRI by exploiting structural similarity," *Inverse Problems*, vol. 31, no. 1, p. 015001, 2014.

[250] M. J. Ehrhardt, P. Markiewicz, M. Liljeroth, A. Barnes, V. Kolehmainen, J. S. Duncan, L. Pizarro, D. Atkinson, B. F. Hutton, S. Ourselin *et al.*, "PET reconstruction with an anatomical MRI prior using parallel level sets," *IEEE transactions on medical imaging*, vol. 35, no. 9, pp. 2189–2199, 2016.

[251] T. Goldstein and S. Osher, "The split Bregman method for L_1-regularized problems," *SIAM journal on imaging sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[252] C. Li, "An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing," Ph.D. dissertation, 2010.

[253] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA journal of numerical analysis*, vol. 8, no. 1, pp. 141–148, 1988.

[254] Y. Wang, W. Yin, and Y. Zhang, "A fast algorithm for image deblurring with total variation regularization," 2007.

[255] B. Wohlberg, "ADMM penalty parameter selection by residual balancing," *arXiv preprint arXiv:1704.06209*, 2017.

[256] C. A. Cocosco, V. Kollokian, R. K.-S. Kwan, G. B. Pike, and A. C. Evans, "Brainweb: Online interface to a 3D MRI simulated brain database," in *NeuroImage*. Citeseer, 1997.

[257] K. Vunckx, A. Atre, K. Baete, A. Reilhac, C. M. Deroose, K. Van Laere, and J. Nuyts, "Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging," *IEEE transactions on medical imaging*, vol. 31, no. 3, pp. 599–612, 2011.

[258] I. B. Malone, R. E. Ansorge, G. B. Williams, P. J. Nestor, T. A. Carpenter, and T. D. Fryer, "Attenuation correction methods suitable for brain imaging with a

PET/MRI scanner: a comparison of tissue atlas and template attenuation map approaches," *Journal of Nuclear Medicine*, vol. 52, no. 7, pp. 1142–1149, 2011.

[259] K. Vunckx and J. Nuyts, "Heuristic modification of an anatomical Markov prior improves its performance," in *IEEE Nuclear Science Symposuim & Medical Imaging Conference.* IEEE, 2010, pp. 3262–3266.

[260] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman *et al.*, "An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest," *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.

[261] V. L. Villemagne, K. E. Pike, G. Chételat, K. A. Ellis, R. S. Mulligan, P. Bourgeat, U. Ackermann, G. Jones, C. Szoeke, O. Salvado *et al.*, "Longitudinal assessment of Aβ and cognition in aging and Alzheimer disease," *Annals of neurology*, vol. 69, no. 1, pp. 181–192, 2011.

[262] Y. Su, T. M. Blazey, A. Z. Snyder, M. E. Raichle, D. S. Marcus, B. M. Ances, R. J. Bateman, N. J. Cairns, P. Aldea, L. Cash *et al.*, "Partial volume correction in quantitative amyloid imaging," *Neuroimage*, vol. 107, pp. 55–64, 2015.

[263] A. Bousse, S. Pedemonte, D. Kazantsev, S. Ourselin, S. Arridge, and B. Hutton, "Weighted MRI-based Bowsher priors for SPECT brain image reconstruction," in *IEEE Nuclear Science Symposuim & Medical Imaging Conference.* IEEE, 2010, pp. 3519–3522.

[264] S. Khohlmyer, C. Stearns, P. Kinahan, and T. Lewellen, "NEMA NU2-2001 performance results for the GE advance PET system," in *2002 IEEE Nuclear Science Symposium Conference Record*, vol. 2. IEEE, 2002, pp. 890–894.

[265] J. Doshi, G. Erus, Y. Ou, S. M. Resnick, R. C. Gur, R. E. Gur, T. D. Satterthwaite, S. Furth, C. Davatzikos, A. N. Initiative *et al.*, "MUSE: MUlti-atlas region Segmentation utilizing Ensembles of registration algorithms and parameters,

and locally optimal atlas selection," *Neuroimage*, vol. 127, pp. 186–195, 2016.

[266] O. G. Rousset, D. L. Collins, A. Rahmim, and D. F. Wong, "Design and implementation of an automated partial volume correction in PET: application to dopamine receptor quantification in the normal human striatum," *Journal of Nuclear Medicine*, vol. 49, no. 7, pp. 1097–1106, 2008.

[267] A. K. Jha, B. Caffo, and E. C. Frey, "A no-gold-standard technique for objective assessment of quantitative nuclear-medicine imaging methods," *Physics in Medicine & Biology*, vol. 61, no. 7, p. 2780, 2016.

[268] A. K. Jha, E. Mena, B. S. Caffo, S. Ashrafinia, A. Rahmim, E. C. Frey, and R. M. Subramaniam, "Practical no-gold-standard evaluation framework for quantitative imaging methods: application to lesion segmentation in positron emission tomography," *Journal of Medical Imaging*, vol. 4, no. 1, p. 011011, 2017.

[269] K. Gong, J. Guan, C.-C. Liu, and J. Qi, "PET image denoising using a deep neural network through fine tuning," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 153–161, 2018.

[270] K. Gong, J. Guan, K. Kim, X. Zhang, J. Yang, Y. Seo, G. El Fakhri, J. Qi, and Q. Li, "Iterative PET image reconstruction using convolutional neural network representation," *IEEE transactions on medical imaging*, vol. 38, no. 3, pp. 675–685, 2018.

[271] K. Kim, D. Wu, K. Gong, J. Dutta, J. H. Kim, Y. D. Son, H. K. Kim, G. El Fakhri, and Q. Li, "Penalized PET reconstruction using deep learning prior and local linear fitting," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1478–1487, 2018.

[272] F. Liu, H. Jang, R. Kijowski, T. Bradshaw, and A. B. McMillan, "Deep learning MR imaging–based attenuation correction for PET/MR imaging," *Radiology*,

vol. 286, no. 2, pp. 676–684, 2018.

[273] T.-A. Song, S. R. Chowdhury, F. Yang, and J. Dutta, "Super-resolution PET imaging using convolutional neural networks," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 518–528, 2020.

[274] ——, "PET image super-resolution using generative adversarial networks," *Neural Networks*, vol. 125, pp. 83–91, 2020.

[275] B. E. Schaafsma, J. S. D. Mieog, M. Hutteman, J. R. Van der Vorst, P. J. Kuppen, C. W. Löwik, J. V. Frangioni, C. J. Van de Velde, and A. L. Vahrmeijer, "The clinical use of indocyanine green as a near-infrared fluorescent contrast agent for image-guided oncologic surgery," *Journal of surgical oncology*, vol. 104, no. 3, pp. 323–332, 2011.

[276] Y. Lou, W. Zhou, T. P. Matthews, C. M. Appleton, and M. A. Anastasio, "Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging," *Journal of biomedical optics*, vol. 22, no. 4, p. 041015, 2017.

Curriculum Vitae

Yansong Zhu was born in Shiyan, China on May 31^{th}, 1993. He received his B.S. degree in Optical Information Science and Technology from Huazhong University of Science and Technology in 2015. He then started his Master program in Electrical and Computer Engineering at the Johns Hopkins University and joined Quantitative Radiomolecular Imaging & Therapy (Qurit) lab under the supervision of Dr. Arman Rahmim. He completed his M.S.E degree in 2016 and started PhD program in Electrical and Computer Engineering at the Johns Hopkins University since 2017. In 2018, Yansong moved to Vancouver, Canada as a visiting graduate student at University of British Columbia. During his PhD, Yansong's research spanned optical imaging using diffuse light as well as positron emission tomography. He was awarded the 2020 Bradley-Alavi fellowship from the Society of Nuclear Medicine & Molecular Imaging (SNMMI) for his work on partial volume correction of brain amyloid PET images.