1-1-2014

Inversion of the Star Transform

Fan Zhao
University of Pennsylvania, zhaofan84@gmail.com

Inversion of the Star Transform

Abstract

We define the star transform as a generalization of the broken ray transform for image reconstruction in single scattering tomography. Using the star transform provides advantages including possibility to reconstruct the absorption and the scattering coefficients of the medium separately and simultaneously. We derive the star transform from physical principles, and derive several computationally efficient algorithms for its inversion. We discuss mathematical properties and analyze numerical stability of inversion, and obtain necessary conditions for stable reconstruction. An approach combining scattered rays and ballistic rays to improve reconstruction is provided, and total variation and L1 regularization are utilized to remove noise. Numerical experiments are carried out to test the algorithms of inversion, the possibility to recover the absorption and the scattering coefficients, and the effect of different regularizations.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Applied Mathematics

First Advisor
Vadim A. Markel

Subject Categories
Applied Mathematics

INVERSION OF THE STAR TRANSFORM

Fan Zhao

A DISSERTATION

in

Applied Mathematics and Computational Science

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2014

Supervisor of Dissertation

_____

Vadim A. Markel, Associate Professor of Radiology and Bioengineering

Graduate Group Chairperson

_____

Charles L. Epstein, Thomas A. Scott Professor of Mathematics, Professor of Radiology in Mathematics

Dissertation Committee:
James C. Gee, Associate Professor of Radiologic Science in Radiology
Vadim A. Markel, Associate Professor of Radiology and Bioengineering
Andrew Tsourkas, Associate Professor of Bioengineering

ABSTRACT
INVERSION OF THE STAR TRANSFORM

Fan Zhao
Vadim A. Markel

We define the star transform as a generalization of the broken ray transform for image reconstruction in single scattering tomography. Using the star transform provides advantages including possibility to reconstruct the absorption and the scattering coefficients of the medium separately and simultaneously. We derive the star transform from physical principles, and derive several computationally efficient algorithms for its inversion. We discuss mathematical properties and analyze numerical stability of inversion, and obtain necessary conditions for stable reconstruction. An approach combining scattered rays and ballistic rays to improve reconstruction is provided, and total variation and L^1 regularization are utilized to remove noise. Numerical experiments are carried out to test the algorithms of inversion, the possibility to recover the absorption and the scattering coefficients, and the effect of different regularizations.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In medical imaging, absorption and scattering are two major factors that affect the data. Different models that describe the propagation of photons are developed during last decades. Models need to be carefully chosen according to the problem. The problem is usually described by its scale compared to the mean free path of the photons [30].

The most widely and successfully used technology is computerized tomography (CT). CT reconstructs the attenuation in the medium that X-rays penetrate through, to provide tomographic images. In CT, scattering is neglected in the model, therefore the propagation of photons is described as straight lines. Mathematically, the projection data of CT are line integrals in 2-D plane. The reconstruction procedure can be considered as inversion of an integral transform consisting of the integrals of a function over straight lines. The transform is known as Radon

transform.

CT can provide high-contrast images with speed and accuracy, but it also has its limitations. One main limitation of conventional CT is that it can not be applied to low-absorption high-scattering medium. Even in low-scattering medium, artifacts caused by scattering affect the image of CT. Therefore, techniques are needed to remove artifacts from the reconstruction [18]. Moreover, CT requires 180 degrees of data collection, due to the requirement of inverting Radon transform, which limits the size of the examined object. One more issue of concern is the radiation dose [15]. CT is regarded as a moderate- to high-radiation technique, which might increase risk, especially when repeated tests are needed.

For higher scattering medium, models with scattering need to be taken into account. A mathematical model for describing photon propagation with scattering is radiative transfer equation (RTE). In RTE, radiance is defined as energy flow per unit normal area per unit solid angle per unit time. Green's function is essential for analytically solving RTE, but the solution can only be obtained for simple cases. Frameworks for solving RTE is given for different geometries [1, 5, 11, 23, 24, 32]. In tomography, we need to solve inverse radiative transfer problem. Inverse source problems and inverse boundary-condition problems with RTE are studied in [10, 33]. Solving the equation numerically is computational expensive, because radiance depends on position, direction and time, which contains six different independent variables. The problem would be six-dimensional for three-dimensional tomography.

For strong scattering medium, a widely used approximation for RTE is diffusion approximation. A modality of medical imaging that has been developed is diffusion optical tomography (DOT), which utilize visible light as source instead of X-ray. It is applicable to breast imaging, and can be used to monitor brain function and pathological changes [4, 9, 16]. In diffusion approximation, the propagation of light can be described with a photon diffusion equation. However, the diffusion approximation can only be applied to medium where scattering coefficient is much larger than absorption coefficient, It's not applicable near the source or surface, with highly directed light, or in biological tissue that scatters light strongly in forward direction. In these situations, more complicated approximation may be required. Moreover, the inverse problem to the diffusion equation is ill-posed due to the physics of the propagation.

Reconstruction on diffusion approximation for highly absorbing media with collimated sources are studied in [34, 38]. A generalized model is provided for both steady-state and frequency-domain, which is suggested to give accurate solution to both low and high scattering media, and also have the capacity to predict the transition from the highly anisotropic light distribution near the collimated light source to the isotropic light distribution in the far field. Therefore, it may be applicable to turbid medium with high absorption, and medium close to the source. These studies suggest that generalized methods may be able to overcome some of the limitations of diffusion approximation.

Another approach in DOT is higher-order approximation. P_N approximation is one of these techniques. In P_N approximation, RTE is expended into spherical harmonics and truncated at order N. When $N = 1$, P_1 approximation gives us the photon diffusion equation in diffusion approximation. Higher order approximation models the anisotropy in the radiation more accurately [3, 19–21]. In application, High-order equation is suggested to be able to applied to regions, such as cerebrospinal fluid (CSF), which the first-order equation fails to describe, and it's also suggested to provide possibility to prevent the ill-posedness in solving the parabolic equation in first-order approximation [21].

Coupled model combining RTE and diffusion approximation is proposed to overcome the limitations when utilizing the two approaches alone [35,36]. In this model, RTE is used in domains that assumption of diffusion is invalid, including proximity of the source and boundary, and diffusion approximation is used elsewhere. Numerical results are obtained with finite element method, and show that the coupled model describes photon migration in low-scattering and non-scattering medium better than only using diffusion approximation.

The model describing scattering in this thesis is on another direction from radiative transfer. We focus on a weak scattering medium, and makes a single scattering assumption. The medium should has less scattering than that for diffusion approximation and RTE, and has more scattering than that for CT. This assumption leads to single scattering tomography (SSOT), that overcomes some limitations of

conventional CT. It introduces new geometries that avoid the restriction of object size, and also provides opportunity to recover scattering and absorption coefficients of the medium.

Theoretically, in SSOT, Radon transform can be generalized to circle [7] parabolas [25], and also to a more general class of curves [8]. Compton scattering tomography is designed to recover the local Compton-scattering cross section. The ray is emitted and collected from one side of the object. This geometry does not require 180° access as required in CT, therefore it has the potential to utilize it for internal structure of large objects. In this modality, attenuation due to absorption is neglected and the measurements are related to the weighted integrals of electron density along circular arcs. An analytical solution to the inverse problem is provided, which is analogous to the convolution backprojection algorithm used in conventional CT [28]. A more general family of circular-arc Radon (CAR) transforms is defined in [27]. The inversion of CAR is presented, which is described as rapid and restricted to an adjustable reduced volume in order to meet the need for imaging of smaller objects in non-destructive industrial testing, as well as in medical imaging.

Here, instead of circles or other smooth curves, we are interested in a class of more simple curves : piecewise line segments. In [37], three types of V-line transforms are defined on a curve formed by a pair of half-lines forming the vertical letter V. An analytical inversion is derived by Fourier transform. Filtered backprojection

reconstruction algorithm is also provided for the V-line transform [26].

In [12], broken-ray Radon (BRT) transform is derived by single scattering approximation to the radiative transport equation (RTE) for known scattering coefficient. BRT is defined on curves formed by a pair of half-line segments. BRT can be considered as a generalization of V-line transform, but it does not require the half-line segments to be symmetrical about vertical axis. The inverse problem can be solved by discretizing the integral equations. Numerical experiment shows that the inverse problem is only mild ill-posed. An analytical inversion formula involving second order derivatives is derived in [14], which is analogous to filtered backprojection formula. In [14], a method for separate reconstructions of absorption and scattering coefficients is suggested. Numerical result shows that image quality is relatively low for discontinuous attenuation coefficient, which also confirms the mild ill-posedness of inverting broken-ray Radon transform.

Approaches to invert BRT through only first order derivatives are developed recently. Using the measurement of three detectors and only one source, reconstruction could be done by a method that involves only first order derivatives of the data. This method is local and does not require a complete data set, so that it provides more freedom in source-detector arrangement. The range of the BRT is also described [22].

In this research, we further develop the methods of [12–14]. We focus on the geometry of an open strip wherein angularly-resolved sources and detectors are em-

ployed but no energy resolution or sensitivity is assumed. We also assume scattering occurs in the entire domain therefore the complete data can be collected. In [12], reconstruction for BRT of a single broken-ray is studied. However, the problem is ill-posed and reconstruction without regularization is unstable, therefore high level regularization is needed for that method.

In this thesis, we define star transform that includes more than one broken-rays, which will be discussed further in the next chapter. Analysis shows that it introduces a less ill-posed inverse problem. Numerical experiments also suggest that the reconstruction is more stable than the one with only one broken-ray. Recovering scattering and absorption coefficients simultaneously is also studied in this thesis. The reconstruction is theoretically feasible and the numerical experiments also show stable results. Experiments with more realistic simulated data are needed for future research.

However, the results in our numerical experiments are susceptible to noise. In order to have better results, we can include data of the ballistic rays, which can be collected simultaneously when the scattered data is measured. Without scattering, data of ballistic ray is supposed to have higher intensity and have less noise, therefore we expect the additional data provided by ballistic rays can help reduce the effect of the noise from the data of scattering rays.

The techniques of SSOT mentioned before do not involve ballistic ray in the models. Including ballistic ray introduces a new generalized model of SSOT, which

provides opportunity to apply SSOT for medium with lower scattering. CT requires 180° projection data of ballistic rays to perform reconstruction, but a full collection of ballistic data may be impossible in SSOT models. However, with the limited ballistic data we are able to collect, we can find reconstruction of SSOT that is consistent with it, without significant change to the image reconstruction procedure. In this thesis, we derive an algorithm in Fourier domain to find the reconstruction consistent with the ballistic data, by only a minor adjustment to the inversion algorithm for star transform. We expect including ballistic ray to be an effective improvement to our model.

Other approaches to improve the image quality is utilizing different regularizations. Tikhonov regularization is added during the procedure of reconstruction to overcome the ill-posedness of the inverse problem, but it does not provide good image when noise level is high. We present algorithms to implement the total variation regularization and L^1 regularization. These two regularizations can help remove the noise while preserving the shape of the image.

Chapter 2

The problem

2.1 Geometry

In single scattering tomography, image reconstruction can be performed for 3-D medium by reconstructing 2-D images slice by slice. In each slice, the problem is reduced to a 2-D problem. Therefore, in this thesis we only focus on the problem of 2-D reconstruction. Suppose the domain for imaging is a 2-D slice of a 3-D medium with $X = $ const. The geometry of this slice is an open strip $\mathbb{S} = \{0 < z < L\}$ as shown in Figure 2.1, which is assumed to be infinitely long in direction of Y-axis.

In contrast to conventional CT, which assume that no scattering occurs, we assume that the majority of incident photons undergo single scattering in the medium. To be specific, we require the width L of the strip along the Z-direction to match the scattering property of the medium, $\mu_s L = O(1)$ (μ_s is scattering coefficient).

The reconstruction is designed based on this assumption.

Suppose a photon incident from one side of the strip undergoes a single scattering and its direction changes. The propagation trajectory is a broken line. The power of the incident ray attenuates along the trajectory, which is caused by two major factors : absorption and scattering. The strength of absorption and scattering at location $\mathbf{R}$ is described by absorption coefficient $\mu_a(\mathbf{R})$ and scattering coefficient $\mu_s(\mathbf{R})$. These two coefficients vary inside the medium, so they can be considered as two non-negative functions spatially distributed on $\mathbb{S}$. The attenuation coefficient is the sum of the previous two :

$$\mu(\mathbf{R}) = \mu_a(\mathbf{R}) + \mu_s(\mathbf{R}) \ . \tag{2.1.1}$$

We will focus on reconstructing function $\mu(\mathbf{R})$ first, and recover $\mu_a(\mathbf{R})$ and $\mu_s(\mathbf{R})$ in the next step.

The data for this problem is obtained in the following way: sources that emit photons are aligned on one side of the strip. On the other side of the strip, we have angularly fixed collimated detectors, which only collect signals of rays with certain outgoing directions. Suppose a photon is sent into the medium at the location Y_1 with direction $\hat{\mathbf{u}}_1$, and it's received at location Y_2 with direction $\hat{\mathbf{u}}_2$. Each pair of (Y_1, Y_2) determines a broken line as shown in Figure 2.2. The vertex of the broken line, $\mathbf{R} = (y, z)$, is where scattering occurs. Suppose the directions $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ are fixed, the data can be considered as a function of (Y_1, Y_2). For $\hat{\mathbf{u}}_1 = (u_{1y}, u_{1z})$,

$\hat{\mathbf{u}}_2 = (u_{2y}, u_{2z})$, $\mathbf{R} = (Y, Z)$, the coordinates of $\mathbf{R}$ and (Y_1, Y_2) have the following relation,

$$Y_1 = Y - \frac{u_{1y}}{u_{1z}} Z , \quad Y_2 = Y - \frac{u_{2y}}{u_{2z}} (L - Z) . \tag{2.1.2}$$

Therefore, the data can also be considered as a function of $\mathbf{R}$. We denote the data collected at Y_2 by $W(\mathbf{R})$. The goal of single scattering tomography is to recover μ from W. We assume that scattering coefficient is strictly positive everywhere in the medium: $\mu_s(\mathbf{R}) > \bar{\mu}_s > 0$ for every $\mathbf{R} \in \bar{\mathbb{S}}$. In order to have $W(\mathbf{R})$ exist on the entire domain $\mathbb{S}$, it is important to have this assumption. Otherwise, $W(\mathbf{R})$ does not exist because no scattering occurs at $\mathbf{R}$. Here $\bar{\mu}_s$ is a positive constant indicates the scattering strength on the background. This assumption also implies that $\mu(\mathbf{R}) > \bar{\mu}_s > 0$ in $\mathbb{S}$. W_0 and W have the following relation,

$$W(\mathbf{R}) = W_0 S_{12} \mu_s(\mathbf{R}) \exp\left\{-\left[I_1(\mathbf{R}) + I_2(\mathbf{R})\right]\right\} , \tag{2.1.3a}$$

$$I_k(\mathbf{R}) = \int_0^{\ell_k(\mathbf{R})} \mu\left(\mathbf{R} + \hat{\mathbf{u}}_k \ell\right) d\ell , \quad k = 1, 2 . \tag{2.1.3b}$$

Here $\ell_k(\mathbf{R})$ is the value that $\mathbf{R} + \hat{\mathbf{u}}_k \ell_k(\mathbf{R}) = 0$, or L, depending on the sign of u_{kz}. S_{12} is the phase function describing the geometry, which only depends on $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$. It's determined when the directions are fixed. W_0 and S_{12} are both assumed to be known constants. We define the signal function Φ by taking logarithm of W,

$$\Phi(\mathbf{R}) = -\log\left(\frac{W(\mathbf{R})}{W_0 S_{12} \bar{\mu}_s}\right) = I_1(\mathbf{R}) + I_2(\mathbf{R}) - \log\left[\mu_s(\mathbf{R})/\bar{\mu}_s\right] . \tag{2.1.4}$$

11

$\Phi(\mathbf{R})$ is the integral of μ along a broken line, plus a term involving scattering coefficient $\mu_s(\mathbf{R})$. If μ_s is a known constant in the medium, $\Phi(\mathbf{R})$ only involves integral of μ. However, when μ_s is not constant, the problem involves both functions μ and μ_s. We'd like first remove μ_s from (2.1.4), so we can turn the problem into an inverse integral problem for μ. The method to remove μ_s is to utilize more signals.

Suppose we have K fixed vectors $\hat{\mathbf{u}}_k$, $k = 1, \ldots, K$. Each vector can be an incident direction or an outgoing direction of the rays, or both. Each pair of them defines a collection of data, so we totally have $K(K-1)/2$ data: W_{ij}, $i, j = 1, \ldots,$ K, $i \neq j$. W_{ij} is the data corresponding to $\hat{\mathbf{u}}_i$ and $\hat{\mathbf{u}}_j$. Figure 2.2 (a) shows the case for $K = 3$. The signal function introduced by $W_{ij}(\mathbf{R})$ is:

$$\Phi_{ij}(\mathbf{R}) = -\log\left(\frac{W_{ij}(\mathbf{R})}{W_0 S_{ij} \bar{\mu}_s}\right) = I_i(\mathbf{R}) + I_j(\mathbf{R}) - \log\left[\mu_s(\mathbf{R})/\bar{\mu}_s\right] . \qquad (2.1.5)$$

All signal functions have three terms, the first two of them are line integrals of μ. The third term which contains $\mu_s(\mathbf{R})$ is shared by all signal functions. This term can be canceled by adding up these functions with weighted factors. We choose coefficients c_{ij}, $i, j = 1, \ldots, K$, $i \neq j$, such that $\sum_{i \neq j} c_{ij} = 0$, and define coefficients $s_i = \sum_{j \neq i} c_{ij}$. Adding up the signal functions Φ_{ij} multiplied by c_{ij}, it

gives us a system of equations without scattering coefficient involved :

$$\Phi(\mathbf{R}) = \sum_{i,j} c_{ij} \Phi_{ij}(\mathbf{R})$$

$$= \sum_{i,j} c_{ij} \left[I_i(\mathbf{R}) + I_j(\mathbf{R}) - \log\left(\mu_s(\mathbf{R})/\bar{\mu}_s\right) \right]$$

$$= \sum_i \sum_{j \neq i} c_{ij} I_i(\mathbf{R}) - \sum_{i \neq j} c_{ij} \log\left(\mu_s(\mathbf{R})/\bar{\mu}_s\right)$$

$$= \sum_{i=1}^{K} s_i I_i(\mathbf{R}) \ . \tag{2.1.6}$$

We define (2.1.6) to be star transform with K branches:

$$\Phi(\mathbf{R}) = \sum_{i=1}^{K} s_i I_i(\mathbf{R}) \ , \quad I_k(\mathbf{R}) = \int_0^{\ell_k(\mathbf{R})} \mu\left(\mathbf{R} + \hat{\mathbf{u}}_k \ell\right) d\ell \ . \tag{2.1.7}$$

The star transform can be well defined for function in $L^1(\mathbb{S})$. In the next chapter, we will present methods to invert star transform.
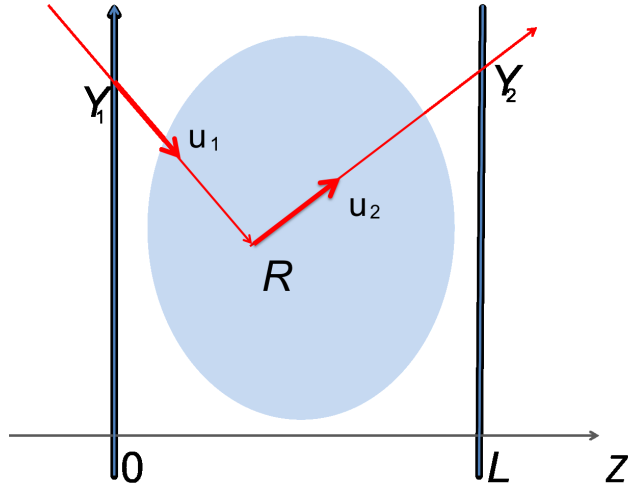
Figure 2.1: A broken line incident at Y_1 in $\hat{\mathbf{u}}_1$, scattering at $\mathbf{R}$, outgoing at Y_2 in $\hat{\mathbf{u}}_2$.

(a)

(b)
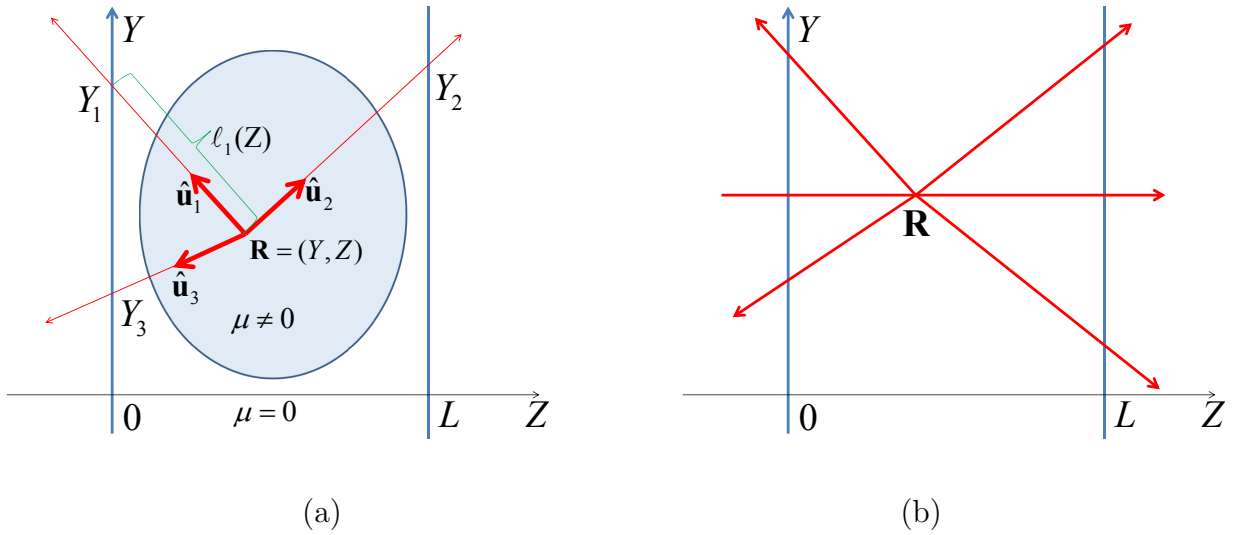
Figure 2.2: (a) Sketch of the imaging geometry for the case $N = 3$ (the distances ℓ_2 and ℓ_3 are not shown). (b) Imaging geometry in which simultaneous measurements of the ballistic and single-scattered rays (due to the same source) are utilized.

Chapter 3

Inversion of the star transform

In this chapter, we turn to the problem of inverting star transform. An simple idea to find the inverse is discretizing the transform. However, attenuation coefficient and the signal function are both 2-D functions, so the computation complexity would be large. Suppose the domain is discretized into $N \times N$ grid, the problem would have N^2 equations of size N^2. This will be infeasible in numerical aspect. Moreover, it's difficult to analyze the stability and add regularization in that approach. In this thesis, we will present an approach to find the inverse in Fourier domain.

3.1   The star transform in Fourier domain

We assume that attenuation coefficient μ inside the medium is compactly supported, so Fourier transforms of Φ and μ are well defined. First, we introduce some notations:

$$\text{Fourier transform:} \qquad \mathscr{F} \ : \ \mathscr{S}(\mathbb{S}) \to \mathscr{S}(\mathbb{S}) \qquad\qquad \text{(3.1.1a)}$$

$$\text{Inverse Fourier transform:} \quad \mathscr{F}^{-1} : \ \mathscr{S}(\mathbb{S}) \to \mathscr{S}(\mathbb{S}) \qquad\qquad \text{(3.1.1b)}$$

$$\text{Star transform:} \qquad \mathcal{F} \ : \ L^1(\mathbb{S}) \to L^1(\mathbb{S}) \qquad\qquad \text{(3.1.1c)}$$

Here $\mathscr{S}(\mathbb{S})$ is the Schwarz space on $\mathbb{S}$, which is a subspace of $L^1(\mathbb{S})$. The reconstruction procedure is

$$\Phi \xrightarrow{\quad \mathscr{F} \quad} \Phi_n(q) \xrightarrow{\quad \text{Inverse problem in Fourier domian} \quad} \mu_n(q) \xrightarrow{\quad \mathscr{F}^{-1} \quad} \mu \ .$$

We first compute Fourier transform of Φ and then solve the inverse problem in Fourier domain. Finally, taking inverse transform gives us the attenuation coefficient μ. The Fourier transform is defined as

$$\mu(y, z) = \int_{-\infty}^{\infty} \frac{dq}{2\pi} e^{iqy} \tilde{\mu}(q, z) = \int_{-\infty}^{\infty} \frac{dq}{2\pi} e^{iqy} \frac{1}{L} \sum_{n=-\infty}^{\infty} \mu_n(q) e^{i\kappa_n z} \ , \qquad \text{(3.1.2a)}$$

$$\mu_n(q) = \int_{-\infty}^{\infty} dy\, e^{-iqy} \int_0^L dz\, e^{-i\kappa_n z} \mu(y, z) \ , \qquad\qquad \text{(3.1.2b)}$$

$$\kappa_n = \frac{2\pi n}{L} \ . \qquad\qquad\qquad\qquad\qquad\qquad \text{(3.1.2c)}$$

In (3.1.2), q is the Fourier parameter corresponding to y, and κ_n is the parameter corresponding to z. Notice that the domain $\mathbb{S}$ is bounded in Z-direction and is

unbounded in Y-direction, this is the reason that the definition of our Fourier transform differs in Y and Z directions. To find the relation between $\Phi_n(q)$ and $\mu_n(q)$, we start by taking Fourier transform (2.1.7) with respect to Y:

$$
\begin{aligned}
\tilde{\Phi}(q, Z) &= \sum_{k=1}^{K} s_k \int_0^{\ell_k(Z)} d\ell \int_{-\infty}^{\infty} \mu(Y + u_{ky}\ell, Z + u_{kz}\ell) e^{iqY} dY \\
&= \sum_{k=1}^{K} s_k \int_0^{\ell_k(Z)} e^{-iqu_{ky}\ell} \tilde{\mu}(q, Z + u_{kz}\ell) d\ell \\
&= \sum_{k=1}^{K} \frac{s_k}{u_{kz}} e^{-i\beta_k(q)Z} \int_Z^{\xi_k} e^{i\beta_k(q)z} \tilde{\mu}(q, z) dz \\
&= \sum_{k=1}^{K} \frac{s_k}{u_{kz}} e^{-i\beta_k(q)Z} \int_Z^{\xi_k} e^{i\beta_k(q)z} \sum_m \mu_m(q) e^{i\kappa_m z} dz \\
&= \sum_{k=1}^{K} \frac{s_k}{u_{kz}} e^{-i\beta_k(q)Z} \sum_{n=-\infty}^{\infty} \frac{e^{i[\beta_k(q)+\kappa_n]\xi_k} - e^{i[\beta_k(q)+\kappa_n]Z}}{i[\beta_k(q) + \kappa_n]} \mu_n(q) .
\end{aligned}
\qquad (3.1.3)
$$

In the above derivation, we have used the fact that the upper limit of integration over ℓ, $\ell_k(Z)$, is independent of Y. The notations in (3.1.3) are defined as

$$
\beta_k(q) = q \frac{u_{ky}}{u_{kz}} , \qquad \xi_k =
\begin{cases}
L , & \text{if } u_{kz} > 0 \\
0 , & \text{if } u_{kz} < 0
\end{cases} .
\qquad (3.1.4)
$$

Here ξ_k is the Z-coordinate of the k-th ray intersection with the boundary of $\mathbb{S}$. Since we assume that all rays intersect the strip boundaries, the quantities in (3.1.4) are well defined. We also have that $\exp(i\xi_k \kappa_n) = 1$ for all k.

To shorten the notations, we will omit the parameter q in (3.1.3) below by writing β_k, μ_n instead of $\beta_k(q)$, $\mu_n(q)$. We then fix q and take Fourier transform

17

of (3.1.3) with respect to Z, which results in the following infinite system of linear equations:

$$\Phi_n = \mu_n \sum_{k=1}^{K} \frac{is_k}{u_{kz}(\beta_k + \kappa_n)} + \sum_{k=1}^{K} \frac{s_k e^{i\beta_k \xi_k} \left(e^{-i\beta_k L} - 1\right)}{Lu_{kz}(\beta_k + \kappa_n)} \sum_{m=-\infty}^{\infty} \frac{\mu_m}{\beta_k + \kappa_m} \; . \qquad (3.1.5)$$

Introducing notations

$$d_n = \sum_{k=1}^{K} \frac{is_k}{u_{kz}(\beta_k + \kappa_n)} = \sum_{k=1}^{K} \frac{is_k}{\hat{\mathbf{u}}_k \cdot (q, \kappa_n)} \; , \qquad \alpha_k = \frac{e^{i\beta_k \xi_k} \left(e^{-i\beta_k L} - 1\right)}{Lu_{kz}} \; , \qquad (3.1.6)$$

we can rewrite (3.1.5) in the form

$$\Phi_n = d_n \mu_n + \sum_{k=1}^{K} \frac{s_k \alpha_k}{\beta_k + \kappa_n} \sum_{m=-\infty}^{\infty} \frac{\mu_m}{\beta_k + \kappa_m} \; . \qquad (3.1.7)$$

Furthermore, (3.1.7) can also be written in matrix form:

$$\mathbf{A}\mathbf{x} = \mathbf{D}\mathbf{x} + \sum_{k=1}^{K} \alpha_k \mathbf{a}_k \mathbf{a}_k^* \mathbf{x} = \boldsymbol{\Phi} \; . \qquad (3.1.8)$$

The matrices and vectors are defined as

$$\mathbf{A} = \mathbf{D} + \sum_{k=1}^{K} \alpha_k \mathbf{a}_k \mathbf{a}_k^* \,, \tag{3.1.9a}$$

$$\mathbf{D} = \begin{pmatrix} \ddots & & & \\ & \sum_{k=1}^{K} \frac{is_k}{u_{kz}(\beta_k + \kappa_n)} & & \\ & & \ddots & \end{pmatrix} \,, \tag{3.1.9b}$$

$$\mathbf{a}_k = \left(\cdots \quad \frac{1}{\beta_k + \kappa_n} \quad \cdots \right)^T \,, \tag{3.1.9c}$$

$$\mathbf{\Phi} = \left(\cdots \Phi_n \cdots \right)^T \,. \tag{3.1.9d}$$

Matrix $\mathbf{A}$ is a diagonal matrix $\mathbf{D}$ with K separable terms (Rank-1 matrix). We can use this feature to find specific algorithm for inversion of $\mathbf{A}$, with advantages in effectiveness and stability.

3.2 Direct inversion

As defined in (3.1.7), each equation has infinitely many terms. Theoretically, an inversion algorithm for (3.1.7) exists. We assume $\mathbf{D}$ in equation (3.1.8) to be invertible. Multiplying the equation from left by $\mathbf{a}_j^* \mathbf{D}^{-1}$, gives us

$$y_j + \sum_{k=1}^{K} M_{jk} y_k = z_j \,, \quad j = 1, \ldots, K \,, \tag{3.2.1}$$

where y_j, M_{jk} and z_j are defined as

$$y_j = \mathbf{a}_j{}^*\mathbf{x} \ , \quad M_{jk} = \alpha_k\mathbf{a}_j{}^*\mathbf{D}^{-1}\mathbf{a}_k \ , \quad z_j = \mathbf{a}_j{}^*\mathbf{D}^{-1}\mathbf{\Phi} \ . \tag{3.2.2}$$

Although (3.1.7) has infinitely many equations and unknowns, (3.2.1) has only K equations with K unknowns. Suppose variables in (3.2.2) are well defined, the system of equations is of a small size K, which can be easily solved. Then the solution $\mathbf{x}$ can be computed by

$$\mathbf{x} = \mathbf{D}^{-1}\mathbf{\Phi} - \sum_{k=1}^{K} \alpha_k y_k \mathbf{D}^{-1}\mathbf{a}_k \ . \tag{3.2.3}$$

$\mathbf{D}$ is a diagonal matrix, (3.2.3) can be computed entry by entry. To be more specific, the nth entry in (3.2.3) is

$$x_n = d_n{}^{-1}\Phi_n - \sum_{k=1}^{K} \alpha_k y_k d_n{}^{-1} a_{kn} \ . \tag{3.2.4}$$

One issue needed to be discussed before using (3.2.3) is whether variables in (3.2.2) are well defined. The variables are defined as sum of infinitely terms, so we need to guarantee those sums converge. Let's first consider M_{jk},

$$M_{jk} = \alpha_k \sum_{n=-\infty}^{\infty} \frac{1}{d_n(\beta_j + \kappa_n)(\beta_k + \kappa_n)} \ . \tag{3.2.5}$$

We will expend the terms into Taylor series with respect to $\kappa_n{}^{-1}$. For $n \neq 0$,

20

$\dfrac{1}{\beta_j + \kappa_n}$ can be expended as

$$
\begin{aligned}
\frac{1}{\beta_j + \kappa_n} &= \kappa_n{}^{-1}(\beta_j/\kappa_n + 1)^{-1} \\
&= \kappa_n{}^{-1}\left[1 - \beta_j\kappa_n{}^{-1} + \beta_j{}^2\kappa_n{}^{-2} + O(\kappa_n{}^{-3})\right] \qquad (3.2.6)\\
&= \kappa_n{}^{-1} - \beta_j\kappa_n{}^{-2} + \beta_j{}^2\kappa_n{}^{-3} + O(\kappa_n{}^{-4}) \ .
\end{aligned}
$$

Then the following functions can also be expended into series with respect to $\kappa_n{}^{-1}$,

$$
\frac{1}{(\beta_j + \kappa_n)(\beta_k + \kappa_n)} = \kappa_n{}^{-2} - (\beta_j + \beta_k)\kappa_n{}^{-3} + O(\kappa_n{}^{-4}) \ , \qquad (3.2.7\text{a})
$$

$$
d_n = t_1\kappa_n{}^{-1} - t_2\kappa_n{}^{-2} + O(\kappa_n{}^{-3}) \ , \qquad (3.2.7\text{b})
$$

$$
d_n{}^{-1} = t_1{}^{-1}\kappa_n + t_1{}^{-2}t_2 + O(\kappa_n{}^{-1}) \ , \qquad (3.2.7\text{c})
$$

where the notations are

$$
t_1 = \sum_{k=1}^{K} is_k/u_{kz} \ , \quad t_2 = \sum_{k=1}^{K} is_k\beta_k/u_{kz} \ .
$$

Therefore, taking into account that $\kappa_n = -\kappa_{-n} = 2\pi n/L$, the terms in (3.2.5) with

21

odd n are canceled. t_1, t_2 are independent with n, so

$$
\begin{aligned}
M_{jk} &= \frac{1}{\beta_j \beta_k d_0} + \sum_{n \neq 0} \{ t_1 \kappa_n^{-1} + [t_1^{-2} t_2 - t_1 (\beta_j + \beta_k)] \kappa_n^{-2} + O(\kappa_n^{-3}) \} \\
&= \frac{1}{\beta_j \beta_k d_0} + \sum_{n=1}^{\infty} \{ [t_1^{-2} t_2 - t_1 (\beta_j + \beta_k)] \kappa_n^{-2} + O(\kappa_n^{-4}) \} \ .
\end{aligned}
$$

$$(3.2.8)$$

$\kappa_n \sim O(n)$, so the sum in (3.2.8) converges. M_{jk} are well defined. Next we turn to

$z_j = \mathbf{a}_j^* \mathbf{D}^{-1} \mathbf{\Phi}$. Using the result in (3.2.7), the nth entry of $\mathbf{a}_j^* \mathbf{D}^{-1}$ is

$$
\left(\mathbf{a}_j^* \mathbf{D}^{-1} \right)_n = t_1^{-1} + O(\kappa_n^{-1}) \ . \tag{3.2.9}
$$

$\left(\mathbf{a}_j^* \mathbf{D}^{-1} \right)_n$ is bounded. $\mathbf{\Phi}$ is the Fourier coefficients of signal function, so $\sum_n \Phi_n < \infty$. Therefore $z_j = \mathbf{a}_j^* \mathbf{D}^{-1} \mathbf{\Phi} < \infty$ is also well defined. In this section we have presented a direct method that is theoretically applicable for even infinite system. In practice, we need to truncate the Fourier series to have a finite system. In the next section, we will talk about an approach to compute the numerical inverse of $\mathbf{A}$ with truncation, which is used in the simulation chapter in this thesis.

3.3 Inversion of star transform in Fourier domain

In this section, we will introduce an iterative method to compute the inverse of (3.1.7). Here we assume the data is discretized on a $N \times N$ grid. Then the matrix $\mathbf{A}$ is $N \times N$, and the vectors in (3.1.9) are $N \times 1$. Before considering inverting (3.1.7), we first start with inverting a matrix that contains only one separable term.

Let this matrix $\mathbf{A}$ be

$$\mathbf{A} = \mathbf{D} + \mathbf{b}\mathbf{a}^* . \tag{3.3.1}$$

$\mathbf{D}$ is a $N \times N$ diagonal matrix, $\mathbf{a}$ and $\mathbf{b}$ are both $N \times 1$ vectors. The inversion formula is derived from Sherman-Morrison formula. We first assume $\mathbf{D}$ to be invertible, to be specific, all diagonal entries of $\mathbf{D}$ are not zero. We also assume $\mathscr{D}$ to be nonzero, which is defined as

$$\mathscr{D} = 1 + \mathbf{a}^* \mathbf{D}^{-1} \mathbf{b} . \tag{3.3.2}$$

When the two assumptions hold, the inversion formula is given by

$$T = \mathscr{D}^{-1}\mathbf{b}\mathbf{a}^* , \quad \mathbf{A}^{-1} = \mathbf{D}^{-1} - \mathbf{D}^{-1}T\mathbf{D}^{-1} . \tag{3.3.3}$$

Requiring $|\mathscr{D}| > 0$ is essential because the determinant of A is given by

$$|\mathbf{A}| = \mathscr{D}\,|\mathbf{D}| = \mathscr{D}\prod_{n=1}^{N} d_n . \tag{3.3.4}$$

So the two assumptions above are actually equivalent to $\mathbf{A}$ is invertible. When utilizing this formula in numerical computation, a small singular value of $\mathbf{D}$ or a small $|\mathscr{D}|$ can lead to unstable result. Therefore, we need more strict conditions. We require $|\mathscr{D}| > \sigma$, where $\sigma > 0$ is some pre-selected small constant. Formula (3.3.3) is only applicable to invertible $\mathbf{A}$, but in some case even $\mathbf{A}$ is not invertible, we can still get a inversion formula. Suppose $\mathbf{D}$ has only one small singular value

ϵ, the inverse of $\mathbf{D}$ is

$$\mathbf{D}^{-1} = \mathbf{D}_{reg}{}^{-1} + \frac{1}{\epsilon}\mathbf{v}\mathbf{w}^* , \quad \mathbf{v}^*\mathbf{D}_{reg} = \mathbf{D}_{reg}\mathbf{w} = 0 . \tag{3.3.5}$$

$\mathbf{v}$ and $\mathbf{w}$ are the singular vectors corresponding to singular value $1/\epsilon$ in the singular value decomposition (SVD) of $\mathbf{D}^{-1}$. Plugging (3.3.5) to (3.3.3), and taking $\epsilon \to 0$, it gives us a formula for inverting $\mathbf{A}$,

$$\mathbf{A}^{-1} = \mathbf{D}_{reg} + \frac{\mathscr{D}_{reg}}{\mathbf{a}^*\mathbf{v}\mathbf{b}^*\mathbf{w}}\mathbf{v}\mathbf{w}^* - \frac{1}{\mathbf{a}^*\mathbf{v}}\mathbf{v}\mathbf{a}^*\mathbf{D}_{reg}{}^{-1} - \mathbf{D}_{reg}{}^{-1}\mathbf{b}\mathbf{w}^*\frac{1}{\mathbf{w}^*\mathbf{b}} , \tag{3.3.6a}$$

$$\mathscr{D}_{reg} = 1 + \mathbf{a}^*\mathbf{D}_{reg}{}^{-1}\mathbf{b} . \tag{3.3.6b}$$

We can use (3.3.3) when $\mathbf{D}$ is invertible, or (3.3.6) when it has a small singular value. In both cases, the inversion for matrix with one separable term is simple to compute. Notice that (3.3.3) and (3.3.6) only involve the inverse of $\mathbf{D}$, so they are applicable as long as we know the inverse of $\mathbf{D}$, even $\mathbf{D}$ is not diagonal.

We can generalize this algorithm for the matrix with K separable terms, by repeating (3.3.3). Suppose we have a matrix

$$\mathbf{A} = \mathbf{D} + \sum_{i=1}^{K} \mathbf{b}_i \mathbf{a}_i^* . \tag{3.3.7}$$

Here $\mathbf{D}$ is also a diagonal matrix and $\mathbf{a}_i$, $\mathbf{b}_i$ are vectors in $\mathbb{C}^N$. Compared with

(3.1.8), we have $\mathbf{b}_i = \alpha_i \mathbf{a}_i$. We will introduce more notations. Let $\mathbf{A}_k$ be the matrix with the first k separable terms, $k \leq K$,

$$\mathbf{A}_k = \mathbf{D} + \sum_{i=1}^{k} \mathbf{b}_i \mathbf{a}_i^* , \quad k = 0, \cdots, K . \tag{3.3.8}$$

The two special cases are $\mathbf{A}_0 = \mathbf{D}$ and $\mathbf{A}_K = \mathbf{A}$. We start with inverting $\mathbf{A}_0$ and iteratively calculate the inverse of $\mathbf{A}_k$ for $k = 1, 2, \ldots, K$. When $\mathbf{A}_k^{-1}$ is obtained, we use (3.3.3) to compute the inverse of $\mathbf{A}_{k+1}$. This is a direct method, because we are guaranteed to find the inverse of $\mathbf{A}$ when the algorithm terminates after K steps. The iteration in detail is

- *First iteration step.* Assume that $\mathbf{D}$ has no small singular value or only one small singular value, otherwise the algorithm is not applicable. Then start with $\mathbf{A}_0 = \mathbf{D}$ and consider the first step of iteration, $\mathbf{A}_1 = \mathbf{A}_0 + \mathbf{b}_1 \mathbf{a}_1^*$.

- *Iteration step.* Assume that $\mathbf{A}_k^{-1}$ is obtained for some $k < K$. Compute the determinant

$$\mathscr{D}_k = 1 + \mathbf{a}_{k+1}^* \mathbf{A}_k^{-1} \mathbf{b}_{k+1}. \tag{3.3.9}$$

If $|\mathscr{D}_k| < \sigma$, where σ is a pre-selected small constant, iteration is theoretically feasible, but numerically unstable. Otherwise, compute the k-th update of the matrix according to

$$T_k = \mathscr{D}_k^{-1} \mathbf{b}_{k+1} \mathbf{a}_{k+1}^*. \tag{3.3.10}$$

25

Next, compute $\mathbf{A}_{k+1}{}^{-1}$ according to

$$\mathbf{A}_{k+1}^{-1} = \mathbf{A}_k^{-1} - \mathbf{A}_k^{-1} T_k \mathbf{A}_k^{-1} \ , \tag{3.3.11}$$

and move to the next iteration step.

- *Termination.* The procedure is terminated after K iterations and converges to $\mathbf{A}^{-1} = \mathbf{A}_K{}^{-1}$ as long as all determinants $\mathscr{D}_k$ $(k = 1, 2, \ldots, K - 1)$ satisfy the condition $|\mathscr{D}_k| > \sigma$.

For this algorithm to be feasible, we assume that all $|\mathscr{D}_k| > 0$. Similar to the case for $K = 1$, $|\mathscr{D}_k| > 0$ is equivalent to $\mathbf{A}_k$ is invertible. Therefore we have a more strict requirement for $\mathbf{A}$: not only $\mathbf{A}$ needs to be invertible but also all the intermediate $\mathbf{A}_k$.

3.4 The special case for $q = 0$

In the last section, we omit a fact: when $q = 0$, all β_k in (3.1.4) are zero, so $\mathbf{D}$, $\mathbf{b}_k$, $\mathbf{a}_k$ are not well defined because all entries with $q = 0$ are infinite. Therefore we need to consider the inversion for the special case $q = 0$ separately. We start with

the derivation of (3.1.3). The equations that we obtain are different from (3.1.7),

$$
\begin{aligned}
\tilde{\Phi}(0, Z) &= \sum_{k=1}^{K} s_k \int_0^{\ell_k(Z)} d\ell \int_{-\infty}^{\infty} \mu(Y + u_{ky}\ell, Z + u_{kz}\ell) dY \\
&= \sum_{k=1}^{K} s_k \int_0^{\ell_k(Z)} \tilde{\mu}(0, Z + u_{kz}\ell) d\ell \\
&= \sum_{k=1}^{K} \frac{s_k}{u_{kz}} \int_Z^{\xi_k} \tilde{\mu}(0, z) dz \\
&= \sum_{k=1}^{K} \frac{s_k}{u_{kz}} \left(\sum_{n \neq 0} \frac{1 - e^{i\kappa_n Z}}{i\kappa_n} \mu_n(0) + (\xi_n - Z)\mu_0(0) \right) .
\end{aligned}
\tag{3.4.1}
$$

Applying Fourier transform with respect to Z to $\tilde{\Phi}(0, Z)$, and taking $e^{i\kappa_n L} = 1$ into account, equation (3.1.7) becomes

$$
\Phi_n = i\Sigma_1 \frac{\mu_n - \mu_0}{\kappa_n} , \quad n \neq 0 ,
\tag{3.4.2a}
$$

$$
\Phi_0 = \frac{L\Sigma_0}{2}\mu_0 - i \sum_{m \neq 0} \frac{\Sigma_1}{\kappa_m} \mu_m ,
\tag{3.4.2b}
$$

where the constants Σ_0 and Σ_1 are

$$
\Sigma_0 = \sum_{k=1}^{K} \frac{s_k}{|u_{kz}|} , \quad \Sigma_1 = \sum_{k=1}^{K} \frac{s_k}{u_{kz}} .
\tag{3.4.3}
$$

27

(3.4.2) is much simpler than (3.1.7). Suppose $|\Sigma_0| > 0$ and $|\Sigma_1| > 0$, (3.4.2) can be solved with

$$\mu_0 = \frac{2}{L\Sigma_0} \sum_m \Phi_m \; , \tag{3.4.4a}$$

$$\mu_n = \mu_0 - i\frac{\kappa_n \Phi_n}{\Sigma_1} \; . \tag{3.4.4b}$$

This algorithm is not only applicable for truncated equations, but also for the direct method with infinite system of equations. In the later one, the term $\sum_m \Phi_m$ equals to $\tilde{\Phi}(0,0)$. The value of Σ_0 and Σ_1 depend on the directions of $\hat{\mathbf{u}}_k$ and the coefficients s_k. We need to prevent them to be too small when choosing $\hat{\mathbf{u}}_k$.

3.5 Pseudo-inverse of star transform in Fourier domain

In the last section, we have discussed an algorithm for inverting (3.1.7). The algorithm works when $\mathbf{A}_k$ are invertible and $\mathscr{D}_k$ are not too small, for $k = 1, 2, \ldots, K$. However, for star transform we are not guaranteed to have them hold. When the requirements are not met, we still want to reconstruct μ. We can compute the pseudo inverse and add regularization. The pseudo inverse of a matrix $\mathbf{A}$ is defined as

$$\mathbf{A}^\dagger = (\mathbf{A}^* \mathbf{A} + \lambda \mathbf{I})^{-1} \mathbf{A}^* \; . \tag{3.5.1}$$

28

Pseudo inverse of a matrix is widely used in solving linear least-square problem, or finding the minimum norm solution to s linear system. For pseudo inverse, we don't need $\mathbf{A}$ to be invertible or even square. Here λ is the regularization parameter. When $\mathbf{A}$ is invertible and $\lambda \to 0$, the pseudo inverse converges to the inverse $\mathbf{A}^{-1}$, so pseudo inverse is an extension of conventional inverse. $\mathbf{A}^{\dagger}\mathbf{x}$ is the minimizer of the optimization problem

$$\min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{\Phi}\|^2 + \lambda \|\mathbf{x}\|^2 \ . \tag{3.5.2}$$

For single scattering tomography, we assume that the number of detectors is no less than the image resolution, so we have on less equations than unknowns. It means that the matrix has more rows than columns. For a $N \times M$ matrix $\mathbf{A}$ $(N \geq M)$, we define two matrices to help compute pseudo-inverse. One is of size $N \times N$ and the other is $M \times M$. The two matrices are defined as

$$\mathbf{S}_N = (\mathbf{A}\mathbf{A}^* + \lambda \mathbf{I}_N)^{-1} \ , \tag{3.5.3a}$$

$$\mathbf{S}_M = (\mathbf{A}^*\mathbf{A} + \lambda \mathbf{I}_M)^{-1} \ . \tag{3.5.3b}$$

Here $\mathbf{I}_N$ and $\mathbf{I}_M$ are the identity matrices of the size $N \times N$ and $M \times M$ respectively. $\mathbf{S}_N$ and $\mathbf{S}_M$ are well defined when $\lambda > 0$. Then the pseudo inverse of $\mathbf{A}$ is given by $\mathbf{A}^{\dagger} = \mathbf{S}_M\mathbf{A}^* = \mathbf{A}^*\mathbf{S}_N$. We will introduce an algorithm similar to the one in Section 3.3. In contrast to the previous one, we need to update both $\mathbf{S}_N$ and $\mathbf{S}_M$ in each

step. We first introduce some notations:

$$\mathbf{S}_{N,k} = (\mathbf{A}_k \mathbf{A}_k^* + \lambda \mathbf{I}_N)^{-1} \, , \qquad (3.5.4a)$$

$$\mathbf{S}_{M,k} = (\mathbf{A}_k^* \mathbf{A}_k + \lambda \mathbf{I}_M)^{-1} \, . \qquad (3.5.4b)$$

These matrices correspond to the pseudo inverse of $\mathbf{A}_k$ defined in (3.3.8). The iteration starts with computing $\mathbf{S}_{N,0}$ and $\mathbf{S}_{M,0}$. For $\lambda > 0$, even $\mathbf{D}$ is singular, $\mathbf{S}_{N,0}$ and $\mathbf{S}_{N,0}$ are well defined. In the case of Star transform, $\mathbf{S}_{N,0}$ and $\mathbf{S}_{M,0}$ are diagonal matrices with the diagonal entries

$$s_m = \frac{1}{|d_m|^2 + \lambda} \, , \quad m = 1 \, , \, 2 \, , \, \cdots \, , \, M \, . \qquad (3.5.5)$$

In the case $N > M$, the matrix $\mathbf{S}_{N,0}$ has s_m as the first M diagonal entries and $1/\lambda^2$ as the last $N - M$ entries, while $\mathbf{S}_{M,0}$ is the $M \times M$ minor of $\mathbf{S}_{N,0}$. Compared to the case in computing inverse in Section 3.3, we also need two matrices $\mathbf{T}_{N,k}$ and $\mathbf{T}_{M,k}$, instead of only one $\mathbf{T}_k$. The main iteration step is

$$\mathbf{S}_{N,k+1} = \mathbf{S}_{N,k} - \mathbf{S}_{N,k} \mathbf{T}_{N,k} \mathbf{S}_{N,k} \, , \qquad (3.5.6a)$$

$$\mathbf{S}_{M,k+1} = \mathbf{S}_{M,k} - \mathbf{S}_{M,k} \mathbf{T}_{M,k} \mathbf{S}_{M,k} \, , \qquad (3.5.6b)$$

$$\mathbf{A}_{k+1}^\dagger = \mathbf{A}_{k+1}^* \mathbf{S}_{N,k+1} = \mathbf{S}_{M,k+1} \mathbf{A}_{k+1}^* \, . \qquad (3.5.6c)$$

Here $\mathbf{T}_{N,k}$ and $\mathbf{T}_{M,k}$ are updated with

$$\mathbf{T}_{M,k} = \frac{1}{\mathscr{D}_k}\Big[\gamma_k \mathbf{a}_{k+1}\mathbf{b}_{k+1}{}^*\mathbf{A}_k + \text{H.c.}$$

$$+ \lambda^2 Q_k \mathbf{a}_{k+1}\mathbf{a}_{k+1}{}^* - P_k \mathbf{A}_k{}^*\mathbf{b}_{k+1}\mathbf{b}_{k+1}{}^*\mathbf{A}_k\Big] \ , \tag{3.5.7a}$$

$$\mathbf{T}_{N,k} = \frac{1}{\mathscr{D}_k}\Big[\gamma_k \mathbf{A}_k \mathbf{a}_{k+1}\mathbf{b}_{k+1}{}^* + \text{H.c.}$$

$$+ \lambda^2 P_k \mathbf{b}_{k+1}\mathbf{b}_{k+1}{}^* - Q_k \mathbf{A}_k \mathbf{a}_{k+1}\mathbf{a}_{k+1}{}^*\mathbf{A}_k{}^*\Big] \ , \tag{3.5.7b}$$

where "H.c." stands for Hermitian conjugate of the preceding term and

$$\gamma_k = 1 + \mathbf{a}_{k+1}{}^*\mathbf{A}_k{}^\dagger\mathbf{b}_{k+1} \ , \tag{3.5.8a}$$

$$P_k = \mathbf{a}_{k+1}{}^*\mathbf{S}_{M,k}\mathbf{a}_{k+1} > 0 \ , \quad Q_k = \mathbf{b}_{k+1}{}^*\mathbf{S}_{N,k}\mathbf{b}_{k+1} > 0 \ , \tag{3.5.8b}$$

$$\mathscr{D}_k = |\gamma_k|^2 + \lambda^2 P_k Q_k > 0 \ . \tag{3.5.8c}$$

Here $\mathbf{a}_k$ is $M \times 1$ vector and $\mathbf{b}_k$ is $N \times 1$ vector. This iteration terminates when $k = K$, and $\mathbf{A}^\dagger = \mathbf{A}\mathbf{S}_{N,K} = \mathbf{S}_{M,K}\mathbf{A}$. Computing the pseudo inverse is more complicated than computing the inverse. The basic requirement is $\mathscr{D}_k > 0$. The analysis of complexity and stability is needed to be studied carefully and it will be discussed in the next chapter.

Chapter 4

Analysis of the algorithms

4.1 Computational complexity

One advantage of the two algorithms in the last chapter is computational effectiveness. Suppose the domain $\mathbb{S}$ is discretized into a $N \times N$ grid. For each system of equations in form (3.1.7), the size of $\mathbf{A}$ is $N \times N$. When computing the inverse $\mathbf{A}^{-1}$, using the algorithm mentioned above is faster than Gaussian Elimination, which takes only $O(N^3)$ computation. The first iteration step is computing $\mathbf{A}_0{}^{-1}$, which is simple because $\mathbf{A}_0 = \mathbf{D}$ is a diagonal matrix. When updating $\mathbf{A}_{k+1}{}^{-1}$ according to (3.3.11), we don't compute the product of $\mathbf{A}_k{}^{-1} T_k \mathbf{A}_k{}^{-1}$ directly. We first compute $\mathbf{A}_k{}^{-1} \mathbf{b}_{k+1}$ and $\mathbf{A}_k{}^{-1} \mathbf{a}_{k+1}$, then compute $\mathbf{A}_k{}^{-1} T_k \mathbf{A}_k{}^{-1} = \mathscr{D}^{-1}(\mathbf{A}_k{}^{-1} \mathbf{b}_{k+1})(\mathbf{A}_k{}^{-1} \mathbf{a}_{k+1})^*$. So we only need to compute product between a matrix and a vector, not between matrices. Other computation in the iteration step

can also be broken into product between matrix and vector, therefore the compu-
tational complexity for each iteration step is $O(N^2)$ and the total complexity for
solving each system of equations by this algorithm is $O(KN^2)$. For star transform,
$K \ll N$ (In our simulation experiments, we use K = 3), so the complexity is much
less than $O(N^3)$.

The case for computing pseudo inverse is similar. When computing $\mathbf{S}_{N,k}\mathbf{T}_{N,k}\mathbf{S}_{N,k}$
and $\mathbf{S}_{M,k}\mathbf{T}_{M,k}\mathbf{S}_{M,k}$, we can always use the definitions of $\mathbf{T}_{N,k}$ and $\mathbf{T}_{M,k}$ to break
them into product of vectors, then all computation are between matrices and vec-
tors. Although the algorithm in computing $\mathbf{A}^\dagger$ is more complicated than computing
inverse, but we don't need to compute product between matrices. The complexity
is also $O(KN^2)$ for solving each system of equations.

For the reconstruction, we have N systems of equations, so the total time needed
is $O(KN^3)$. We should also take the computation for Fourier transform into ac-
count. The computational amounts for 2-D Fourier transform and inverse Fourier
transform are $O(N^2 \log N)$ when utilizing Fast Fourier Transform (FFT). Therefore
the total computational amount needed is $O(KN^3)$. For comparison, suppose we
discretize star transform (2.1.7) directly, the number of equations and unknowns
are both N^2, therefore the size of the problem is N^2. Using Gaussian Elimination
to solve the problem takes computation as much as $O(N^6)$. So the complexity for
our algorithm is much less than the direct method.

4.2 Stability

Another issue needed to be considered is the stability of the algorithms. The computing of the inverse is stable only if diagonal matrix $\mathbf{D}$ is stably inverted. For a fixed parameter q, the entry of $\mathbf{D}$ is defined in (3.1.9). Let's write the nth entry $d_n(q)$ in the form of

$$d_n(q) = \sum_{k=1}^{K} \frac{is_k}{\hat{\mathbf{u}}_k \cdot (q, \kappa_n)} = \frac{i}{|(q, \kappa_n)|} \sum_{k=1}^{K} \frac{s_k}{\mathbf{u}_k \cdot \hat{\mathbf{v}}} \ , \tag{4.2.1}$$

where $\hat{\mathbf{v}} = (q, \kappa_n)/\sqrt{q^2 + \kappa_n^2}$ is a unit vector pointing to the direction of (q, κ_n). We can consider $d_n(q)$ as a function of (q, κ_n). If there exists some (q, κ_n) such that $d_n(q) \approx 0$, it leads to instability when inverting $\mathbf{D}$. Define function

$$f(\theta) = \sum_{k=1}^{K} \frac{s_k}{\cos(\theta - \theta_k)} = \sum_{k=1}^{K} \frac{s_k}{\mathbf{u}_k \cdot \hat{\mathbf{v}}} \ , \tag{4.2.2}$$

where θ and θ_k are the angles of $\hat{\mathbf{v}}$ and $\hat{\mathbf{u}}_k$ with the positive direction of Z-axis. $d_n(q) = 0$ for some (q, κ_n) if and only if $f(\theta) = 0$ for some θ. One thing needs to mention is that $f(\theta + \pi) = f(\theta)$, therefore, it is sufficient to consider the interval $0 \leq \theta < \pi$. Suppose $f(\theta)$ has a zero point $\tilde{\theta}$ in $[0, \pi)$, then for (q, κ_n) with angle close to $\tilde{\theta}$, the entry $d_n(q)$ would be very small. We expect $f(\theta)$ to have no zeros, so we can prevent the singularity of $\mathbf{D}$.

We have found necessary conditions for $f(\theta)$ has no zero: (i) First, K needs to be odd; (ii) The directions of all $s_k \hat{\mathbf{u}}_k$ should not be placed on some half plane

in YZ plane. Before we prove that these are necessary conditions, we first clarify some facts about $f(\theta)$. Let $\eta_{k,\pm}$ denote $\theta_k \pm \pi/2$. Notice that $f(\theta) \to \infty$ when θ goes to $\eta_{k,\pm}$ (In geometry, this is when $\hat{\mathbf{v}}$ is perpendicular to $\hat{\mathbf{u}}_k$), so $\eta_{k,\pm}$ are the singular points of $f(\theta)$. Between two adjacent singular points, $f(\theta)$ is a continuous function. Therefore, between two adjacent $\eta_{k,\pm}$, if $f(\theta) \to \infty$ with different signs as θ approaches to the two ends of the interval, there is at least one zero point inside the interval. Another thing worth noticing is the signs of $\dfrac{s_k}{\cos(\theta - \theta_k)}$ are different inside $(\eta_{k,-}, \eta_{k,+})$ and outside this interval. Therefore, for each $\eta = \eta_{k,\pm}$, $f(\theta)$ has different signs in interval $(\eta - \epsilon, \eta)$ and $(\eta, \eta + \epsilon)$, where ϵ is an infinitesimal constant.

With all the observations above, we can show the two conditions mentioned above are necessary. Let's place $s_k \hat{\mathbf{u}}_k$ onto the plane one by one. The first unit vector $s_1 \hat{\mathbf{u}}_1$ can be placed on the plane arbitrarily, and its perpendicular line divides $[0, 2\pi)$ into two intervals as shown in Figure 4.1 (a). In the figure, we use a circle to indicate $[0, 2\pi)$ and the two half circles are the two intervals divided by the perpendicular line of $\hat{\mathbf{u}}_1$. In each of this interval, $f(\theta)$ has the same sign near the boundary of the interval. This is also easy to see from the formula of $f(\theta)$ with $K = 1$. When adding the second vector $s_2 \hat{\mathbf{u}}_2$, suppose $s_2 \hat{\mathbf{u}}_2$ has a different direction with $s_1 \hat{\mathbf{u}}_1$, each of the half circles is divided into two. No matter how we place the second vector, there must be two intervals in which $f(\theta)$ changes sign. This is shown in Figure 4.1 (b). So $f(\theta)$ must have zero points for $K = 2$. Suppose we are placing the third vector, if we place it such that its perpendicular line lies

35

in the interval with different signs, we have a chance to prevent the appearance of the intervals in which $f(\theta)$ has different signs. Figure 4.2 can help us prove this. Suppose we already have two vectors placed in the plane. As discussed above, now $[0, 2\pi)$ is divided into four intervals, in two of which $f(\theta)$ have the same sign near boundary and in the other two $f(\theta)$ have different signs. If we want $f(\theta)$ have the same sign near boundary in all intervals after putting $s_3\hat{\mathbf{u}}_3$, $\theta_3 \pm \pi/2$ must be placed in the intervals in which $f(\theta)$ changes sign, like the case in Figure 4.2 (a). In this case, the angles between any two adjacent $s_k\hat{\mathbf{u}}_k$ are less then $\pi/2$. Otherwise, as shown in Figure 4.2 (b), there must be an interval in which $f(\theta)$ changes sign, so $f(\theta)$ have zero points in this interval. Figure 4.3 is the plot of $f(\theta)$ corresponding to Figure 4.2, we can see when conditions (ii) is not satisfied, $f(\theta)$ has zeros. This necessary condition can be generalized for $K > 3$ in the same way.

However, even conditions (i) and (ii) are satisfied, it does not guarantee that $f(\theta)$ has no zero. This is because even $f(\theta)$ has the same sign near the two ends of a interval, it can still cross the line $f(\theta) = 0$ inside the interval. Figure 4.4 shows the case when (i) and (ii) hold but $f(\theta)$ still has zeros. In the simulation section, we will demonstrate that whether $f(\theta)$ has zeros is essential to the stability of inversion. Therefore for more stable reconstruction, we need to choose $\hat{\mathbf{u}}_k$ and s_k to make sure $|f(\theta)| > \sigma > 0$, where σ is some small constant.

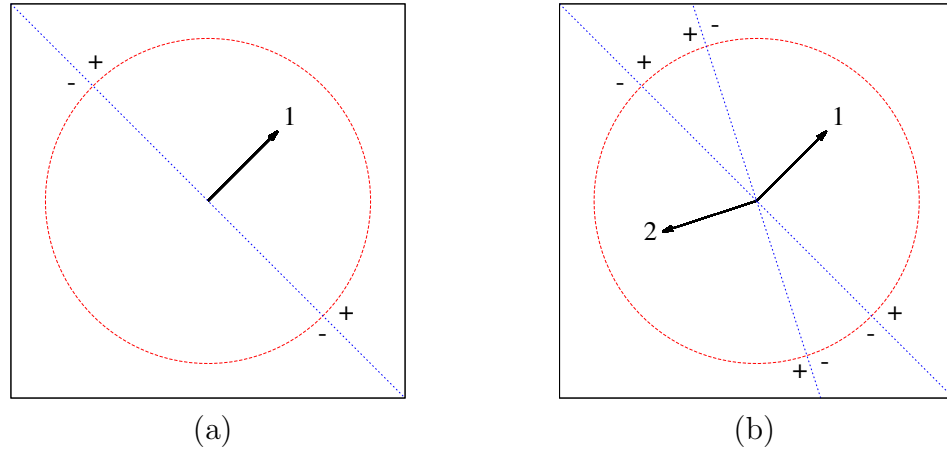Now we turn to the updating steps in the algorithm. By (3.3.10), if we can

36

(a)   (b)

Figure 4.1: (a) Placing $s_1\hat{\mathbf{u}}_1$, the unit circle is divided into two parts. (b) Placing the second vector $s_2\hat{\mathbf{u}}_2$, the unit circle is divided into four parts. The labels $+$ or $-$ indicates the sign of $f(\theta)$ when θ approaches the blue dash line. In Figure (b), $f(\theta)$ has zeros in the top-right and bottom-left intervals.

(a)   (b)

Figure 4.2: Sign diagrams for the function $f(\theta)$ in the case $K = 3$ and $\theta_1 = -\pi/6$, $\theta_2 = 0.82\pi$, $\theta_3 = 0.23\pi$. Angles are measured from the positive direction of the Z-axis, which runs horizontally from left to right. The coefficients s_k are $s_1 = s_2 = 1$, $s_3 = 2$ (a) and $s_1 = s_2 = 1$, $s_3 = -2$ (b). It can be seen that the drawing of the third line, which goes in this case from the top left to the bottom right corner, removes the contradiction in the case (b) but not in (a). By inspection, it can be verified that in (a) all vectors $s_k\hat{\mathbf{u}}_k$ can be placed in the same half-plane while in (b) the same is not true.
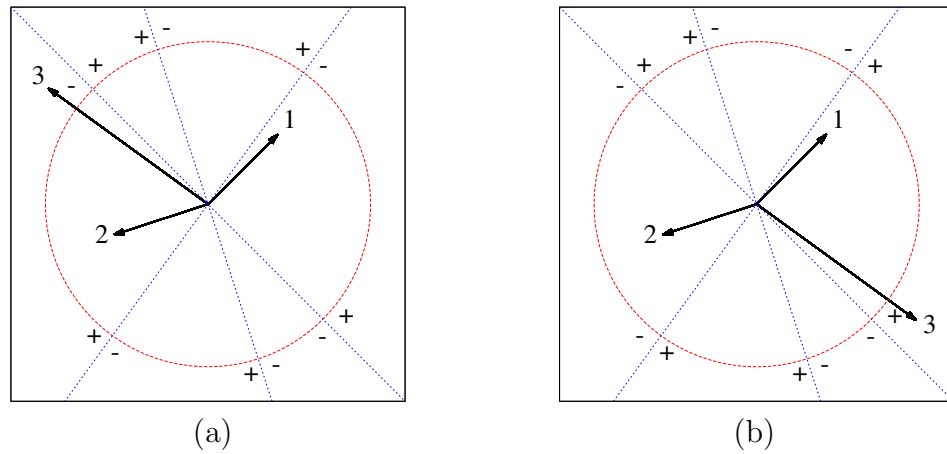
37

(a)   (b)

Figure 4.3: Plot for $f(\theta)$ in the case $K = 3$ and $\theta_1 = -\pi/6$, $\theta_2 = 0.82\pi$, $\theta_3 = 0.23\pi$. Angles are measured from the positive direction of the Z-axis, which runs horizontally from left to right. The coefficients s_k are $s_1 = s_2 = 1$, $s_3 = 2$ (a) and $s_1 = s_2 = 1$, $s_3 = -2$ (b). $f(\theta)$ has zeros in (a) but no zero in (b).
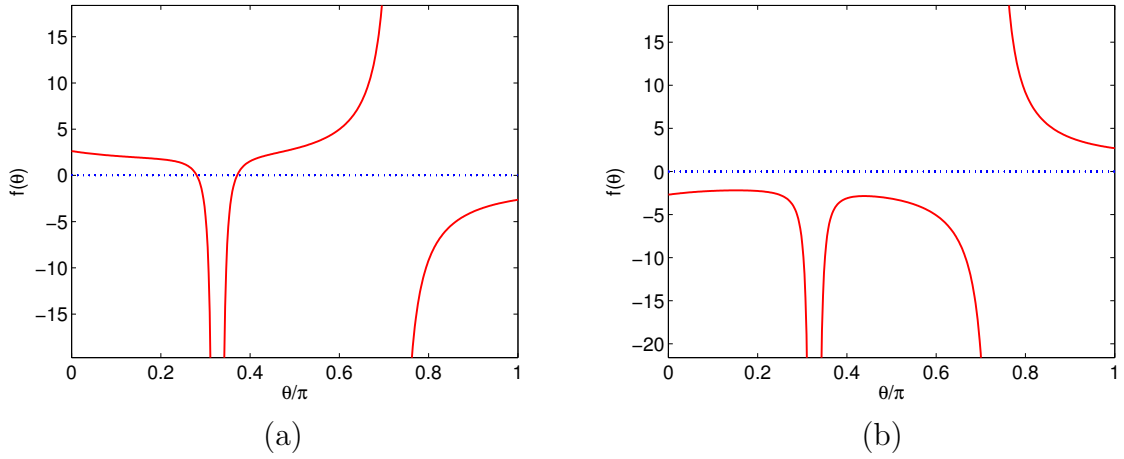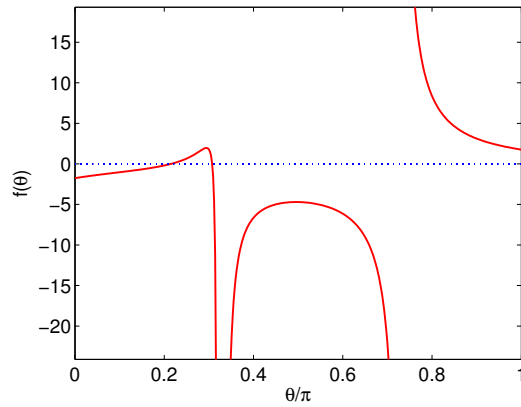

Figure 4.4: Plot for $f(\theta)$ in the case $K = 3$ and $\theta_1 = -\pi/6$, $\theta_2 = 0.82\pi$, $\theta_3 = 0.23\pi$ and coefficients are $s_1 = 1.4$, $s_2 = 0.6$, $s_3 = 2$. $\hat{\mathbf{u}}_k$ satisfy necessary conditions (i) and (ii) but $f(\theta)$ still has zeros.

guarantee $\mathscr{D}_k$ not to be too small, the process will be stable. We have discussed before, $\mathscr{D}_k$ is the determinant of $\mathbf{A}_k$, therefore the stability of reconstruction depends on the invertibility of all $\mathbf{A}_k$.

In the computation of pseudo inverse, adding regularization can help reduce the instability. In each iteration step, since we have $\lambda > 0$, $\mathbf{S}_{M,k}$ and $\mathbf{S}_{N,k}$ in (3.5.4) are both positive definite. Hence P_k and Q_k in (3.5.8) are strictly larger than 0, so $\mathscr{D}_k > 0$ for each step k. The computing the pseudo-inverse is stable.

To be more specific, suppose $\mathbf{A}_k = \mathbf{U\Sigma V}^*$ is the singular value decomposition (SVD) of $\mathbf{A}_k$, we have $\mathbf{S}_{M,k} = \mathbf{V}(\mathbf{\Sigma}^2 + \lambda^2 \mathbf{I}_M)\mathbf{V}^*$. Similarly, $\mathbf{S}_{N,k} = \mathbf{U}(\mathbf{\Sigma}^2 + \lambda^2 \mathbf{I}_N)\mathbf{U}^*$. Then P_k and Q_k can be written as

$$P_k = \sum_{i=1}^{M} \frac{|(\mathbf{Va}_k)_i|^2}{\sigma_i^2 + \lambda^2} \; , \quad Q_k = \sum_{i=1}^{N} \frac{|(\mathbf{Ub}_k)_i|^2}{\sigma_i^2 + \lambda^2} \; . \tag{4.2.3}$$

Here $(\mathbf{Va}_k)_i$ is the ith entry of $\mathbf{Va}_k$, $(\mathbf{Ub}_k)_i$ is the ith entry of $\mathbf{Ub}_k$, and σ_i is the ith diagonal entry of $\mathbf{\Sigma}$. In Q_k, $\sigma_i = 0$ for $i > M$. Then the second term in (3.5.8c) is

$$\lambda^2 P_k Q_k = \lambda^2 \left(\sum_{i=1}^{M} \frac{|(\mathbf{Va}_k)_i|^2}{\sigma_i^2 + \lambda^2} \right) \left(\sum_{i=1}^{N} \frac{|(\mathbf{Ub}_k)_i|^2}{\sigma_i^2 + \lambda^2} \right) \; . \tag{4.2.4}$$

Suppose $\mathbf{A}_k$ is singular, $|\gamma_k| = 0$. In this case, there is some $\sigma_n = 0$, so $\mathscr{D}_k \sim O(\lambda^{-2})$. If we choose regularization parameter λ properly, we can guarantee $\mathscr{D}_k$ not to be too small.

Chapter 5

Reconstruction of scattering and absorption coefficients

In the last two chapters, we have discussed the method for reconstructing atten-
uation coefficient μ, now we turn to reconstruction of scattering and absorption
coefficients. As mentioned before, one advantage of single scattering tomography
over CT is that it provides an opportunity to recover the scattering coefficient and
also absorption coefficient. This is because in our model scattering coefficient $\mu_s(\mathbf{R})$
is a weighted factor in signal function $\Phi_{jk}(\mathbf{R})$. We rewrite signal function (2.1.5) as

$$\mu_s(\mathbf{R}) = \bar{\mu}_s \exp\left[I_j(\mathbf{R}) + I_k(\mathbf{R}) - \Phi_{jk}(\mathbf{R})\right] \ . \tag{5.0.1}$$

Here j and k can be any two distinct numbers from 1, 2, $\ldots$, K. Since $\Phi_{jk}(\mathbf{R})$ is
data and $\bar{\mu}_s$ is assumed to be known, the only unknown terms on the right hand side

are $I_j(\mathbf{R})$ and $I_k(\mathbf{R})$. This two terms are the line integrals of μ along broken line with vertex $\mathbf{R}$. One direct way to compute $I_j(\mathbf{R})$ and $I_k(\mathbf{R})$ is numerical integral. In this thesis, we compute them in Fourier domain as well, and compute the sum $I_j + I_k$ together instead of computing them separately. Let

$$\psi(\mathbf{R}) = I_j(\mathbf{R}) + I_k(\mathbf{R}) \ . \tag{5.0.2}$$

The Fourier coefficients of ψ is defined as

$$\psi_n(q) = \int_{-\infty}^{\infty} dy e^{-iqy} \int_0^L dz e^{-i\kappa_n z} \psi(y, z) \ . \tag{5.0.3}$$

Use the same method in deriving (3.1.7), the Fourier coefficients for ψ has the relation with the Fourier coefficients of μ,

$$\psi_n = d_n \mu_n + \frac{\alpha_j}{\beta_j + \kappa_n} \sum_m \frac{\mu_m}{\beta_j + \kappa_m} + \frac{\alpha_k}{\beta_k + \kappa_n} \sum_m \frac{\mu_m}{\beta_k + \kappa_m} \ , \tag{5.0.4}$$

Here β_j, β_k are defined in (3.1.4), and κ_n is the parameter defined in (3.1.2). The matrices are the same as defined in Section 3.1. We also omit the q in the same way in Section 3.1. This is actually a direct problem for the broken-ray Radon transform. The Fourier coefficients of μ can be obtained during the reconstruction of μ. After computing Fourier coefficients of ψ by (5.0.4), we take inverse Fourier

41

transform with FFT, then we will have $\psi = I_j + I_k$:

$$\mu_n(q) \xrightarrow{\quad (5.0.4) \quad} \psi_n(q) \xrightarrow{\quad \mathscr{F}^{-1} \quad} \psi \ .$$

Then $\mu_s(\mathbf{R})$ can be computed by (5.0.1). We could use any signal function Φ_{jk} to compute μ_s, or we could use more than one Φ_{jk} and compute the average of the results. After recovering μ_s, the absorption coefficients μ_a is the difference between μ and μ_s:

$$\mu_a(\mathbf{R}) = \mu(\mathbf{R}) - \mu_s(\mathbf{R}). \qquad (5.0.5)$$

In contrast to the reconstruction of μ, reconstructions of μ_s and μ_a do not involve solving inverse integral problem. The procedure mentioned above won't cause instability. However, the reconstructions of μ_s and μ_a highly depend on the reconstruction of μ.

For the aspect of computational complexity, computing Fourier coefficients of $\psi = I_j + I_k$ takes $O(N^2)$, and computing Fourier transform costs $O(N^2 \log N)$. This approach for computing ψ takes totally $O(N^2 \log N)$. It just takes a little more computation for recovering scattering and absorption coefficients. This is another advantage of single scattering tomography.

Chapter 6

Improving the image quality

Our imaging reconstruction assumes single scattering. If this assumption holds, it provides us a chance to recover the scattering and absorption coefficients. However, the scattering phenomenon also causes significant loss of the power of incident rays. It also causes the signal to be susceptible to noise. Simulation results show that the reconstruction suffers badly from high level noise. In this section, we will talk about some approaches to reduce the effect of noise and improve the image quality.

6.1   Including ballistic data

The first approach is to include some less noised data to correct the image affected by noise. In SSOT, ballistic ray are not included in the theory, but it exists and has stronger intensity with less noise. Adding ballistic ray to our model may provide better reconstruction. In CT, reconstructing μ with only ballistic rays requires 180-

degree projection data. Without significant change to our model, we can collect ballistic data for at most K directions, so we can not recover μ from only ballistic data. In my thesis, we expect to find reconstruction that is consistent with ballistic data. In this section, we will discuss an algorithm to reconstruct μ with both scattered rays and ballistic rays, which only requires minor adjustments to data collection and inversion algorithm introduced in Chapter 2 and 3.

We can collect the ballistic data at the same time of collecting scattering data. Collecting ballistic data requires an additional collimated detector. While some of incident photons undergo scattering inside the medium, a large quantity of photons do not. These photons travel with straight trajectory, so detectors on the other side of the strip facing the direction of incident rays can collect the projection data of the ballistic rays. Like the signal in conventional CT, the power of the ballistic rays attenuates and it also contains information of the attenuation coefficients μ. Without scattering, the signal of the ballistic rays has larger intensity. We expect including the ballistic ray can help improve the imaging quality.

Suppose an incident ray is sent into medium in direction $\hat{\mathbf{u}}_j$ at Y, the power detected at the opposite side of the strip is

$$W_j(Y) = W_0 \exp[-I_j(Y)] \;, \quad I_j(Y) = \int_0^{\ell_j(Y)} \mu(Y + \hat{u}_{jy}\ell, \hat{u}_{jz}\ell)d\ell \;. \qquad (6.1.1)$$

Here W_0 is the power of the incident ray. For fixed $\hat{\mathbf{u}}_j$, W is a function only defined

44

for one parameter Y. Define the signal function for ballistic data as

$$\Psi_j(Y) = -\log\left[W_j(Y)/W_0\right] = \int_0^{\ell_j(Y)} \mu(Y + \hat{u}_{jy}\ell, \hat{u}_{jz}\ell)d\ell \; . \tag{6.1.2}$$

This signal function is the line integral of μ on the straight line that photons travels through. Taking Fourier transform, we have the equation in Fourier domain

$$\tilde{\Psi}_j(q) = \sum_n \mu_n(q)\frac{e^{i(\beta_j(q)+\kappa_n(q))/L} - 1}{i\hat{u}_{jz}(\beta_j(q) + \kappa_n(q))} \; . \tag{6.1.3}$$

The parameters β and κ_n are defined the same as in Section 3.1. To simplify the notations, we omit q and use vector $\mathbf{c}_j$ to denote the coefficients in the right hand side of equation (6.1.3) :

$$\mathbf{c}_j = \left(\cdots, \frac{e^{i(\beta_j+\kappa_n)/L} - 1}{i\hat{u}_{jz}(\beta_j + \kappa_n)}, \cdots\right)^* \; . \tag{6.1.4}$$

Suppose we are inverting star transform with K branches and ballistic data for $\hat{\mathbf{u}}_j$, $j = 1, 2, \ldots, R$ $(R \leq K)$. For each q, combining the equations from star transform and ballistic rays, we have system of equations:

$$\mathbf{Dx} + \sum_{k=1}^{K} \mathbf{b}_k\mathbf{a}_k^*\mathbf{x} = \Phi \; , \tag{6.1.5a}$$

$$\mathbf{c}_j^*\mathbf{x} = \tilde{\Psi}_j \; , \quad j = 1, 2, \ldots, R \; . \tag{6.1.5b}$$

45

This is an overdetermined system of equations. We can use the extra equations as constraints. As the same as before, we use $\mathbf{A}$ to denote $\mathbf{D} + \sum_{k=1}^{K} \mathbf{b}_k \mathbf{a}_k^*$. We introduce the optimization problem with constraints:

$$\min_{\mathbf{x}} \quad \|\mathbf{A}\mathbf{x} - \mathbf{\Phi}\|^2$$
$$\text{s. t.} \quad \mathbf{c}_j^* \mathbf{x} = \tilde{\Psi}_j \ , \quad j = 1, 2, \ldots, R \ . \tag{6.1.6}$$

This is a constrained linear optimization problem. The solution to this problem satisfies the Karush-Kuhn-Tucker conditions (KKT conditions):

$$\mathbf{A}^*\mathbf{A}\mathbf{x} - \mathbf{A}^*\mathbf{\Phi} + \sum_{j=1}^{R} \lambda_j \mathbf{c}_j = 0 \ , \tag{6.1.7a}$$

$$\mathbf{c}_j^* \mathbf{x} = \tilde{\Psi}_j \ , \quad j = 1, \ldots, R \ . \tag{6.1.7b}$$

The vector $\Lambda = (\ldots, \lambda_j, \ldots)$ here is the KKT multipliers, not the regularization parameter mentioned in the previous sections. The objective of the optimization function is convex and the constraints are linear, so the KKT conditions are sufficient and necessary conditions for the solution. We have an approach to find the analytic solution to (6.1.6). To derive the approach, we first consider another optimization problem without constraint:

$$\min_{\mathbf{x}} L(\mathbf{x}) = \|\mathbf{A}\mathbf{x} - \mathbf{\Phi}\|^2 + \tau \sum_{j=1}^{R} \left\| \mathbf{c}_j^* \mathbf{x} - \tilde{\Psi}_j \right\|^2 \ . \tag{6.1.8}$$

46

This is a least square problem, for which we can find the analytic solution. The second term is a penalty term and τ is the penalty parameter. When $\tau \to \infty$, the solution of this problem converges to the solution to (6.1.6). Assume $\mathbf{A}^*\mathbf{A} + \tau \sum_{j=1}^{R} \mathbf{c}_j\mathbf{c}_j^*$ is not singular, the solution of (6.1.8) is unique and it's the zero point of the derivative of $L(\mathbf{x})$. The derivative of the objective function $L(\mathbf{x})$ is

$$\partial_{\mathbf{x}} L(\mathbf{x}) = 2\mathbf{A}^*\mathbf{A}\mathbf{x} - 2\mathbf{A}^*\mathbf{\Phi} + 2\tau \left[\sum_{j=1}^{R} \mathbf{c}_j\mathbf{c}_j^*\mathbf{x} - \tilde{\Psi}_j\mathbf{c}_j \right] . \tag{6.1.9}$$

The solution to (6.1.8) is the solution to $\partial_{\mathbf{x}} L(\mathbf{x}) = 0$. The solution is given by

$$\bar{\mathbf{x}} = \left[\mathbf{A}^*\mathbf{A} + \tau \sum_{j=1}^{R} \mathbf{c}_j\mathbf{c}_j^* \right]^{-1} \left(\mathbf{A}^*\mathbf{\Phi} + \tau \sum_{j=1}^{R} \tilde{\Psi}_j\mathbf{c}_j \right) . \tag{6.1.10}$$

Before taking $\tau \to \infty$, we can simplify the expression of (6.1.10). The idea is to apply Sherman-Morrison formula to $[\mathbf{A}^*\mathbf{A} + \tau \sum_{j=1}^{R} \mathbf{c}_j\mathbf{c}_j^*]^{-1}$. Suppose $\mathbf{A}^*\mathbf{A}$ is invertible, we use $\mathbf{S}$ to denote $(\mathbf{A}^*\mathbf{A})^{-1}$. We first consider the simplest case of (6.1.8) with $R = 1$. The matrix $[\mathbf{A}^*\mathbf{A} + \tau\mathbf{c}_1\mathbf{c}_1^*]^{-1}$ is simplified as $\mathbf{S} - \dfrac{\tau\mathbf{S}\mathbf{c}_1\mathbf{c}_1^*\mathbf{S}}{1 + \tau\mathbf{c}_1^*\mathbf{S}\mathbf{c}_1}$. So the solution for $R = 1$ is

$$\bar{\mathbf{x}} = \mathbf{S}\mathbf{A}^*\mathbf{\Phi} - \frac{\tau(\mathbf{c}_1^*\mathbf{S}\mathbf{A}^*\mathbf{\Phi} - \tilde{\Psi}_1)}{1 + \tau\mathbf{c}_1^*\mathbf{S}\mathbf{c}_1}\mathbf{S}\mathbf{c}_1 . \tag{6.1.11}$$

Based on the assumption $\mathbf{A}$ is invertible, $\mathbf{S}$ is positive definite, so $\mathbf{c}_1^*\mathbf{S}\mathbf{c}_1 > 0$.

Taking $\tau \to \infty$, (6.1.11) turns to be

$$\bar{\mathbf{x}} = \mathbf{S}\mathbf{A}^*\mathbf{\Phi} - \frac{\mathbf{c}_1{}^*\mathbf{S}\mathbf{A}^*\mathbf{\Phi} - \tilde{\Psi}_1}{\mathbf{c}_1{}^*\mathbf{S}\mathbf{c}_1}\mathbf{S}\mathbf{c}_1 \ . \tag{6.1.12}$$

It's easy to check (6.1.12) satisfies the KKT conditions (6.1.7), with $\lambda_1 = -(\mathbf{c}_1{}^*\mathbf{S}\mathbf{A}^*\mathbf{\Phi} - \tilde{\Psi}_1)/\mathbf{c}_1{}^*\mathbf{S}\mathbf{c}_1$. The first term on the right hand side of (6.1.12) is the solution to the inverse star transform problem in Section 3.3. Therefore, this solution can be considered as a projection of our previous solution to the linear space $\{\mathbf{x} \in \mathbb{C}^M | \mathbf{c}_1{}^*\mathbf{x} = \tilde{\Psi}_1\}$.

This can be generalized for $R > 1$ by applying Sherman-Morrison formula for R times. Define a R-column matrix with the constraint vectors $\mathbf{c}_j$, $j = 1, \ldots, R$:

$$\mathbf{C} = [\mathbf{c}_1, \cdots, \mathbf{c}_R] \ , \tag{6.1.13}$$

and a $R \times R$ matrix

$$\mathbf{H} = \mathbf{C}^*\mathbf{S}\mathbf{C} = \begin{bmatrix} \mathbf{c}_1{}^* \\ \ldots \\ \mathbf{c}_R{}^* \end{bmatrix} \mathbf{S} \left[\mathbf{c}_1, \cdots, \mathbf{c}_R\right] \ . \tag{6.1.14}$$

Assume that $\mathbf{H}$ is invertible, the expression (6.1.10) can be written as

$$\bar{\mathbf{x}} = \mathbf{S}\mathbf{A}^*\boldsymbol{\Phi} - \tau\mathbf{S}\mathbf{C}\left(\mathbf{I}_R + \tau\mathbf{H}\right)^{-1}\left(\mathbf{C}^*\mathbf{S}\mathbf{A}^*\boldsymbol{\Phi} - \boldsymbol{\Psi}\right) \ . \qquad (6.1.15)$$

Taking $\tau \to \infty$, (6.1.15) goes to

$$\bar{\mathbf{x}} = \mathbf{S}\mathbf{A}^*\boldsymbol{\Phi} - \mathbf{S}\mathbf{C}\mathbf{H}^{-1}\left(\mathbf{C}^*\mathbf{S}\mathbf{A}^*\boldsymbol{\Phi} - \boldsymbol{\Psi}\right) \ . \qquad (6.1.16)$$

(6.1.16) also satisfies the KKT (6.1.7) for $R > 1$, with $\Lambda = \mathbf{H}^{-1}\left(\mathbf{C}^*\mathbf{S}\mathbf{A}^*\boldsymbol{\Phi} - \boldsymbol{\Psi}\right)$. The discussion above is for invertible $\mathbf{A}$. In the case $\mathbf{A}$ is not invertible or even not square, the derivation is the same. We only need to replace $\mathbf{S}$ by $\left(\mathbf{A}^*\mathbf{A} + \lambda\mathbf{I}_M\right)^{-1}$. Here λ is the regularization parameter defined in Section 3.5.

When applying this formula to numerical experiments, we first compute the first term following the procedure in Section 3.5, then compute the second term. The computational complexity for the second term by formula (6.1.16) is $O(N^2 R)$. In our numerical experiments, $R = 1$ or 2, therefore it is $O(N^2)$.

6.2 Total variation regularization

In the last section, we discussed a possible approach to remove noise, which uses ballistic data to correct the image. In this section, we will talk about another method to improve image quality without extra information. We'd like to utilize

the total variation regularization to help reduce the noise. Total variation denoising is widely used in image processing. It's supposed to have advantage in preserving edges of the objects while removing noise. We will talk about how to utilize total variation denoising during reconstruction.

Suppose f is a differentiable function defined on $\mathbb{S} \subset \mathbb{R}^2$, and its derivative is integrable, the total variation of f is defined as

$$\mathrm{TV}(f) = \int_{\mathbb{S}} |\nabla f(y,z)| dy dz \ , \quad \nabla f(y,z) = (\partial_y f(y,z), \partial_z f(y,z)) \ . \qquad (6.2.1)$$

It's L^1 norm of the derivative of f. For $\phi \in \mathcal{C}_c^1(\mathbb{S})$, by Green's identity,

$$\int_{\mathbb{S}} f \, div\phi = - \int_{\mathbb{S}} \phi \, div f \ . \qquad (6.2.2)$$

The definition can be extended to function $f \in L^1(\mathbb{S}^2)$ by (6.2.2),

$$\mathrm{TV}(f) = \sup \left\{ \int_{\mathbb{S}} f \, div\phi \ : \ \phi \in \mathcal{C}_c^1(\mathbb{S}) \ , \|\phi\|_{L^\infty} \leq 1 \right\} \ . \qquad (6.2.3)$$

For 2-D image F, we can also define the total variation. For $N \times M$ image $\{F_{i,j}\}$, $i = 1, 2, \ldots, N$, $j = 1, 2, \ldots, M$, the total variation is

$$\mathrm{TV}(F) = \sum_{i,j} \sqrt{|F_{i+1,j} - F_{i,j}|^2 + |F_{i,j+1} - F_{i,j}|^2} \ . \qquad (6.2.4)$$

We assume attenuation coefficient is total variation bounded, i.e. $\mu \in \mathcal{B} = \{f :$

TV(f) < ∞}. We will use total variation as a penalty term to form an optimization problem. Unlike the optimization problem mentioned in previous section, this optimization is defined for the attenuation coefficient μ, instead of the Fourier coefficients of μ. The optimization problem is

$$\min_{\mu} \; L(\mu) = \|\mathcal{F}\mu - \Phi\|_2^2 + \tau \text{TV}(\mu) \; . \tag{6.2.5}$$

Here $\mathcal{F}$ is the star transform, Φ is 2-D signal function, and τ is the regularization parameter to control the level of regularization. The TV term is not differentiable, so this optimization problem can't be solved directly by gradient descent method or other gradient dependent methods. In this thesis, we use Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) to minimize $L(\mu)$.

6.2.1 Brief description of FISTA

Fast Iterative Shrinkage-Thresholding Algorithm (FISTA) is a class of iterative methods for solving inverse problems in image processing [2]. The general FISTA is designed for problem

$$\min \; F(x) \equiv f(x) + g(x) \; , \tag{6.2.6}$$

where $g : \mathbb{R}^n \to \mathbb{R}$ is a continuous convex function, which is possibly nonsmooth. $f : \mathbb{R}^n \to \mathbb{R}$ is a smooth convex function which is continuous differentiable with

Lipschitz continuous gradient L. Lipschitz constant L is a number such that

$$\|\nabla f(x) - \nabla f(y)\| \leq L\|x - y\| , \quad \forall \; x, \; y \; \in \mathbb{R}^n . \tag{6.2.7}$$

In [2], a minimizer

$$P_L(y) = \arg\min\left\{g(x) + \frac{L}{2}\left\|x - \left(y - \frac{1}{L}\nabla f(y)\right)\right\|^2\right\} . \tag{6.2.8}$$

is defined. FISTA with constant step size starts with initial $y_1 = x_0$ and $t_1 = 1$. The main iteration step updates x_k and y_k with

$$x_k = P_L(y_k) , \tag{6.2.9a}$$

$$t_{k+1} = \frac{1 + \sqrt{1 + 4t_k^2}}{2} , \tag{6.2.9b}$$

$$y_{k+1} = x_k + \left(\frac{t_k - 1}{t_{k+1}}\right)(x_k - x_{k-1}) . \tag{6.2.9c}$$

It has been proved that x_k converges to the optimal point $\bar{x}$ with rate

$$F(x_k) - F(\bar{x}) \leq \frac{L\|x_0 - \bar{x}\|^2}{(k+1)^2} . \tag{6.2.10}$$

The main task in applying FISTA is to compute $P_L(x)$. When we turn to our problem, the function $f(\mu) = \|\mathcal{F}\mu - \Phi\|_2^2$, and $g(\mu) = \tau\text{TV}(\mu)$. Both of them satisfy the requirements in FISTA. Our f is the L^2 norm of a linear operator of μ, the

gradient $\nabla f(\mu) = 2\mathcal{F}^*\mathcal{F}\mu - 2\mathcal{F}^*\Phi$. Here $\mathcal{F}^*$ is used to denote the conjugate operator of $\mathcal{F}$ with respect to the inner product of $L^2(\mathbb{S})$: $\langle f, g \rangle_2 = \int_{\mathbb{S}} f(y, z)^* g(y, z) dy dz$. The minimizer (6.2.8) becomes

$$P_L(\mu) = \arg\min_x \left\{ \tau \mathrm{TV}(x) + \frac{L}{2} \left\| x - \left(\mu - \frac{2}{L}(\mathcal{F}^*\mathcal{F}\mu - \mathcal{F}^*\Phi) \right) \right\|^2 \right\} . \quad (6.2.11)$$

By Parseval's theorem, for any $f, \; g \in L^2(\mathbb{S})$, we have

$$\langle \mathcal{F}f, g \rangle_2 = \langle \mathscr{F}\mathbf{A}\mathscr{F}^{-1}f, g \rangle_2 = \langle f, \mathscr{F}\mathbf{A}^*\mathscr{F}^{-1}g \rangle_2 . \quad (6.2.12)$$

Here $\mathbf{A}$ is the operator defined in Fourier domain corresponding to star transform. So the conjugate $\mathcal{F}^*$ is given by $\mathscr{F}\mathbf{A}^*\mathscr{F}^{-1}$. In other word, computing $\mathcal{F}^*$ is similar to compute the direct problem of star transform. The only difference is we need to replace $\mathbf{A}(q)$ by $\mathbf{A}(q)^*$ for each q in Fourier domain. When computing $\mathcal{F}^*\mathcal{F}\mu$, we can do it in the same way, only replacing $\mathbf{A}(q)$ by $\mathbf{A}(q)^*\mathbf{A}(q)$.

For our problem, we can choose $L = 2\|\mathcal{F}^*\mathcal{F}\|_2 = 2\sup\{\|\mathcal{F}^*\mathcal{F}x\|_2 : \|x\|_2 = 1\}$. To be specific, by Parseval's theorem,

$$\|\mathcal{F}^*\mathcal{F}\|_2 = \max_q \|\mathbf{A}(q)\|_2^2 .$$

Therefore the Lipschitz constant $L = 2\max_q \|\mathbf{A}(q)\|_2^2$.

53

6.2.2 Computing P_L

Computing P_L is to solve the optimization problem (6.2.8). This problem is different from the original problem (6.2.5), because it removes the linear operator $\mathcal{F}$ from x. Let

$$y = \frac{2}{L}(\mathcal{F}^*\mathcal{F}\mu - \mathcal{F}^*\Phi) \,, \qquad (6.2.13)$$

the minimization problem of solving P_L is a general total variation denoising problem for y. We use an algorithm in [6] to solve this problem. The algorithm is designed for minimizing the total variation of image, and the convergence has also been proved. It's used to solve problem in the form

$$\min_x \frac{\|x - y\|^2}{2\lambda} + \mathrm{TV}(x) \,. \qquad (6.2.14)$$

In [6], the solution is given by $x = y - \pi_{\lambda\mathbf{K}}(y)$. Here $\pi_{\lambda\mathbf{K}}$ is a projection operator to closed convex set

$$\mathbf{K} = \{div\ p\ :\ p \in \mathbf{Y}\ ,|p_{i,j}| \leq 1\ ,\forall\ i,\ j = 1,\ldots,N\} \,. \qquad (6.2.15)$$

The p^n in the algorithm is in $\mathbf{Y} = \mathbb{R}^{N \times N} \times \mathbb{R}^{N \times N}$. $p_{i,j}$ is the (i,j) element of p. The $divp$ is defined as

$$(divp)_{ij} = \begin{cases} p_{i,j}^1 - p_{i-1,j}^1 & 1 < i < N \\ p_{i,j}^1 & i = 1 \\ -p_{i-1,j}^1 & i = N \end{cases} + \begin{cases} p_{i,j}^2 - p_{i,j-1}^2 & 1 < j < N \\ p_{i,j}^2 & j = 1 \\ -p_{i-1,j}^2 & j = N \end{cases} . \qquad (6.2.16)$$

$\pi_{\lambda\mathbf{K}}$ can be computed by a semi-implicit gradient descent algorithm. For pre-fixed parameter $\sigma > 0$, choosing initial $p^0 = 0$, for step $n \geq 0$, the iteration is given by

$$p_{i,j}^{n+1} = \frac{p_{i,j}^n + \sigma(\nabla(divp^n - y/\lambda))_{i,j}}{1 + \sigma|(\nabla(divp^n - y/\lambda))_{i,j}|} . \qquad (6.2.17)$$

Reference [6] has shown that $\lambda divp^n \to \pi_\lambda(y)$ when $\sigma \leq 1/8$.

6.2.3   Algorithm for solving optimization problem with TV regularization

Combining the two algorithms above, we have an algorithm to solve (6.2.5). It's an iterative algorithm, in which each step is also an optimization problem. To sum up, the iteration for our problem is

- *Initialize* : Choose initial μ_0. We can use the result in Section 3.3 as initial value;

- *Iterative step* : Updating μ_j by FISTA with constant step size in (6.2.9). The Lipschitz parameter L is given by $2\|\mathcal{F}^*\mathcal{F}\|_2 = 2\max_q \|\mathbf{A}(q)\|_2{}^2$. The minimizer $P_L(\mu_j)$ is computed by solving

$$\min_x \ \frac{\|x - (\mu_j - 2(\mathcal{F}^*\mathcal{F}\mu_j - \mathcal{F}^*\Phi)/L)\|^2}{2\lambda} + \mathrm{TV}(x) \ ,$$

  according to algorithm described in Section 6.2.2. with $\lambda = \tau/L$.

- *Termination* : Terminate iteration with some stopping criterion.

We expect the total variation regularized optimization can help reduce the noise. However, one disadvantage of this approach is that this algorithm is not fast due to the large amount of computation required for the iteration. We can compute $\mathcal{F}^*\mathcal{F}\mu_j$ and $\mathcal{F}^*\Phi$ by direct problem in Fourier domain to reduce the computation.

6.3 L^1 regularization

Another regularization that could be combined to the reconstruction is L^1 regularization. We use L^1 norm as penalty term instead of the total variation mentioned in the last section.

In general, the L^1 regularization problem is looking for the solution to the optimization problem

$$\min_x \ L(x) = \|\mathcal{A}x - \Phi\|_2^2 + \tau\|x\|_1 \ . \tag{6.3.1}$$

This optimization problem can also be solved by FISTA [2]. The iterative step is

$$x_{k+1} = \mathrm{T}_{\lambda t}\left(x_k - 2t\mathcal{A}^T(\mathcal{A}x_k - \mathbf{b})\right) \ . \tag{6.3.2}$$

Here t is the step size, and $\mathrm{T}_{\lambda t}$ is the shrinkage operator defined by

$$\mathrm{T}_{\alpha}(x)_i = \left(|x_i - \alpha|\right)_+ \mathrm{sgn}(x_i) \ . \tag{6.3.3}$$

When applying L^1 regularization to image processing, x in (6.3.1) could be the coefficients of the image with respect to some wavelet base. Here we follow this approach. Suppose $\mathbf{H}$ is the Haar transform, and x is the Haar coefficients of μ, $x = \mathbf{H}\mu$. We construct an optimization problem

$$\min_x \ L(x) = \|\mathcal{F}\mathbf{H}^{-1}x - \Phi\|_2^2 + \tau\|x\|_1 \ , \tag{6.3.4}$$

and the attenuation coefficient $\mu = \mathbf{H}^{-1}x$. Therefore, we can apply FISTA to (6.3.4), with $\mathcal{A} = \mathcal{F}\mathbf{H}^{-1}$.

Chapter 7

Simulations

In this chapter, we will present numerical experiments. Without real data, we use simulation to 'generate' data. Through the experiments, we intend to use our algorithm to compute the inverse and pseudo inverse of star transform, and check the stability of reconstruction. We also have experiments on recovering μ_s and μ_a simultaneously. Finally, we are planning to validate feasibility of methods for improving image quality, which are introduced in Chapter 6.

7.1   Generating data

We simulate data by 'inverse crime', which means we compute the data under single scattering assumption according to (2.1.3). We compute $I_k(\mathbf{R})$ by numerical integral with small step size. The constant S_{ij} in our experiment is chosen to be $1/4\pi$, which indicates isomorphic scattering. W_0 is used to describe the power of

incident rays. We use W_0 as a parameter to control the level of noise in the data.

To check stability of the reconstruction, we add noise to data. The noise is added according to the physical principle of single scattering tomography: Since the intensity of received ray is proportional to the number of photons received, and the number of photons follows the Poisson distribution, we add Poisson noise to the data. W_0 is used to describe the number of incident photons, and W_{ij} is the number of photons received. A large W_0 indicates a low level of noise. When $W_{ij}(\mathbf{R})$ is computed according to (2.1.3), we replace it with $\tilde{W}_{ij}(\mathbf{R}) = Pois(W_{ij}(\mathbf{R}))$, where $Pois$ is probability density of Poisson distribution. The mean and variance of $Pois(W_{ij}(\mathbf{R}))$ are both $W_{ij}(\mathbf{R})$. As W_0 increases, $W_{ij}(\mathbf{R})$ increases, therefore the noised data $\tilde{W}_{ij}(\mathbf{R})$ has larger mean and variance. As $W_0 \to \infty$, $\tilde{W}_{ij}(\mathbf{R})$ converges to $W_{ij}(\mathbf{R})$. The signal data Φ for reconstruction is weighted sum of the noised data $\tilde{W}_{ij}(\mathbf{R})$. In our experiments, noise is added with $W_0 = 1.6 \times 10^5$, 4×10^4, and 10^4.

The domain $\mathbb{S}$ is discretized into grid with step $h = L/125$. Here L is the width of the strip. In Y direction, the strip is truncated and sampled with the same step size h, for $M = 625$. We have samples in Y direction more than in Z direction. In reconstruction, we have $M = 625$ systems of equations to solve in Fourier domain.

7.2 Numerical experiments

7.2.1 Reconstruction of attenuation coefficient

We first present numerical experiments to validate the stability of inversion of star transform. We have experiments for $K = 2$ and $K = 3$. For $K = 2$, we provide two choices of $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ for comparison. They differ from each other in direction. When we are inverting star transform for $K = 2$, the two sets of broken-line signals Φ_{13} and Φ_{23} share the same $\hat{\mathbf{u}}_3$. Therefore, using parameters $s_1 = -s_2 = 1$, the terms with integral in direction $\hat{\mathbf{u}}_3$ are canceled. The specific parameters are listed in case (a) and (b) of Table 7.1. For $K = 3$, we provide two choices of $\hat{\mathbf{u}}_k$ as well. In one of them, the directions of all $s_k \hat{\mathbf{u}}_k$ are in a half plane, while in the other one are not. For all cases, we have $\sum_{k=1}^{K} s_k = 0$. The parameters are listed in case (c) and (d) of Table 7.1. In Table 7.1, θ_k is the angle of $\hat{\mathbf{u}}_k$ with the positive direction of Z-axis. Σ_0, Σ_1 and the number of zeros of $f(\theta)$ in interval $[0, \pi)$ are also listed in the table. In Table 7.1, we also list the coefficients s_k and the weighted factors c_{ij} for Φ_{ij}. In all cases except case (d), $f(\theta)$ have zeros, so we expect more stable reconstruction for case (d). The vectors $s_k \hat{\mathbf{u}}_k$ in each case are shown in Figure 7.1.

We use two phantoms: a square and the Shepp-Logan phantom. The attenuation coefficient varies from $0.005h^{-1}$ to $0.05h^{-1}$. In order to focus on reconstruction of attenuation coefficient μ, the scattering coefficient in this experiment are constant $\bar{\mu}_s = 0.005h^{-1}$. The phantoms are shown in Figure 7.2. The simulated data for

square with parameters in case (d) is shown in Figure 7.3. The signal-noise-rate for the three noised data are 20.6914, 10.3210 and 5.1229. Therefore the noise is very high for $W_0 = 1 \times 10^4$. When computing pseudo inverse, regularizations are also added. The regularization parameter λ is from 10^{-7} to 10^{-1}.

Case	a	b	c	d
K	2		3	
s_1	1		1	
s_2	-1		1	
s_3	0		-2	
c_{12}	N/A		2	
c_{13}	1		-1	
c_{23}	-1		-1	
θ_1/π	0.82	1	0.0	
θ_2/π	0.23	0.25	0.80	
θ_3/π	-0.1		1.25	0.25
Zeros	1	1	2	0
Σ_0	-0.1488	-0.4142	-0.5924	-0.5924
Σ_1	-2.5175	-2.4142	2.5924	-3.0645

Table 7.1: Parameters in simulation: s_k are weighted factors corresponding to $\hat{\mathbf{u}}_k$; c_{ij} are the weighted factors for Φ_{ij}; θ_k are the angles of $\hat{\mathbf{u}}_k$ with the positive direction of Z-axis; The last three rows present the number of zeros of function $f(\theta)$ in $[0, \pi)$, Σ_0 and Σ_1. Σ_0 and Σ_1 are not very small in all cases.

The results without noise are shown in Figure 7.4. All results have artifacts. The appearance of artifacts is due to the discontinuity of the phantoms. The results for Shepp-Logan phantom are worse than those for square phantom in all cases, because Shepp-Logan has more complicated structure than a square. We expect the result in case (d) to be better, since the parameters in (d) satisfies the conditions described in Section 4.2. The result validates our expectation. In contrast, case (a) and (c) suffer from artifacts severely. The result for case (b) also shows very stable
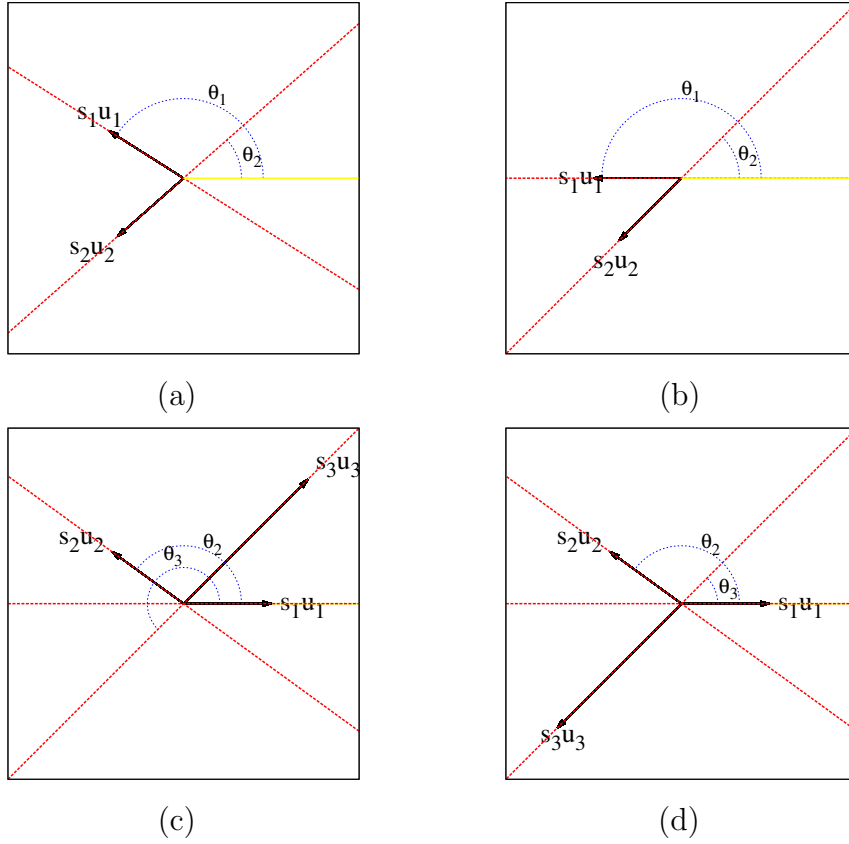
Figure 7.1: Figures (a) - (d) correspond to case (a) - (d) in Table 7.1. Vector labeled $s_k\mathbf{u}_i$ in each figure indicates the directions of $s_k\hat{\mathbf{u}}_i$. In case (c), the directions of all $s_k\hat{\mathbf{u}}_k$ are in one half plane, but in case (d) they are not. θ_k is the angle of $\hat{\mathbf{u}}_k$ with the positive direction of Z-axis.
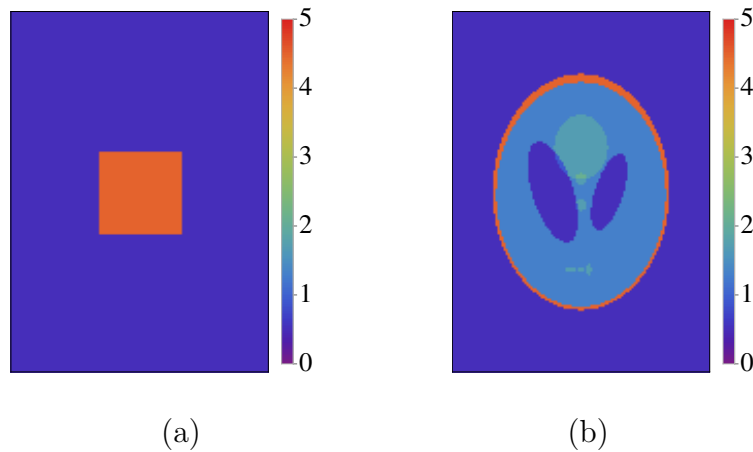

Figure 7.2: (a) The phantom of a square in the center; (b) The Shepp-Logan phantom. The background $\bar{\mu} = \bar{\mu}_s = 0.005h^{-1}$.
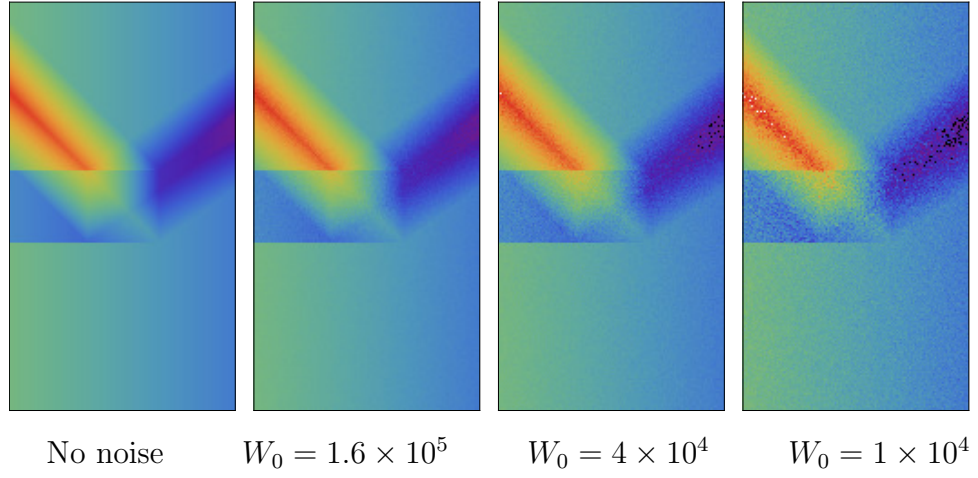
No noise $W_0 = 1.6 \times 10^5$ $W_0 = 4 \times 10^4$ $W_0 = 1 \times 10^4$

Figure 7.3: Simulated data Φ with different noise levels for case (d).

Case a   Case b

Case c   Case d
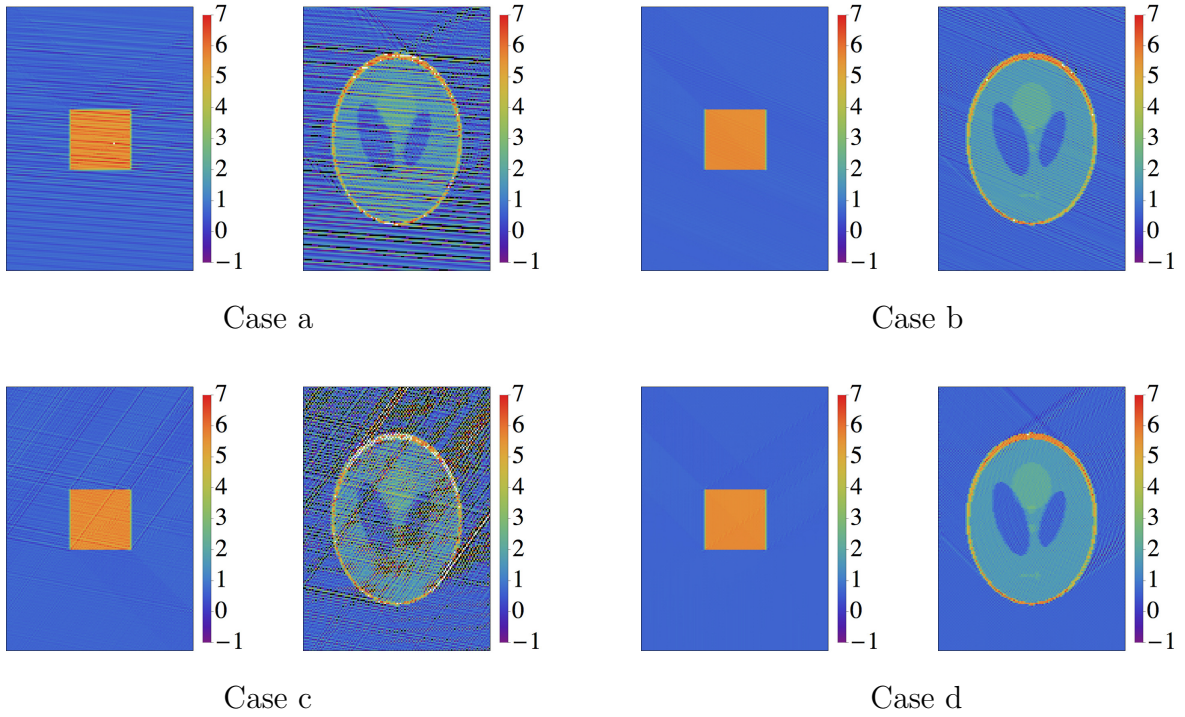
Figure 7.4: Reconstruction with no noise in experiment 1.

Case a

Case b

Case c

Case d
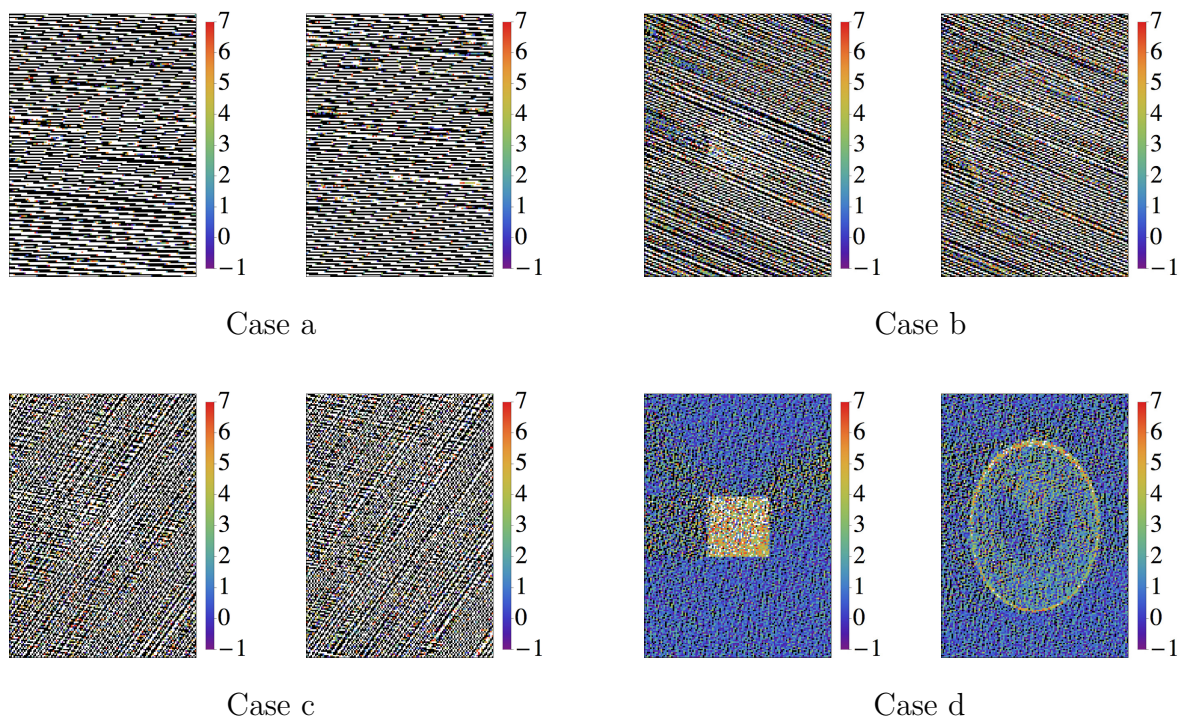
Figure 7.5: Reconstruction without regularization, with noise level $W_0 = 1.6 \times 10^5$ in experiment 1.

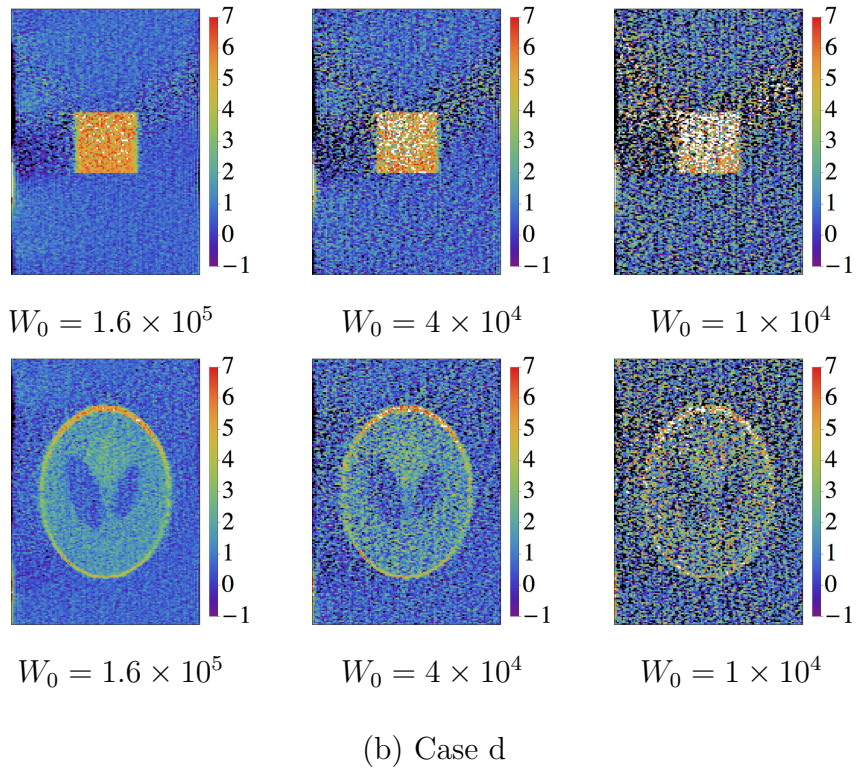$W_0 = 1.6 \times 10^5$ $W_0 = 4 \times 10^4$ $W_0 = 1 \times 10^4$

$W_0 = 1.6 \times 10^5$ $W_0 = 4 \times 10^4$ $W_0 = 1 \times 10^4$

(a) Case c

$W_0 = 1.6 \times 10^5$ $W_0 = 4 \times 10^4$ $W_0 = 1 \times 10^4$

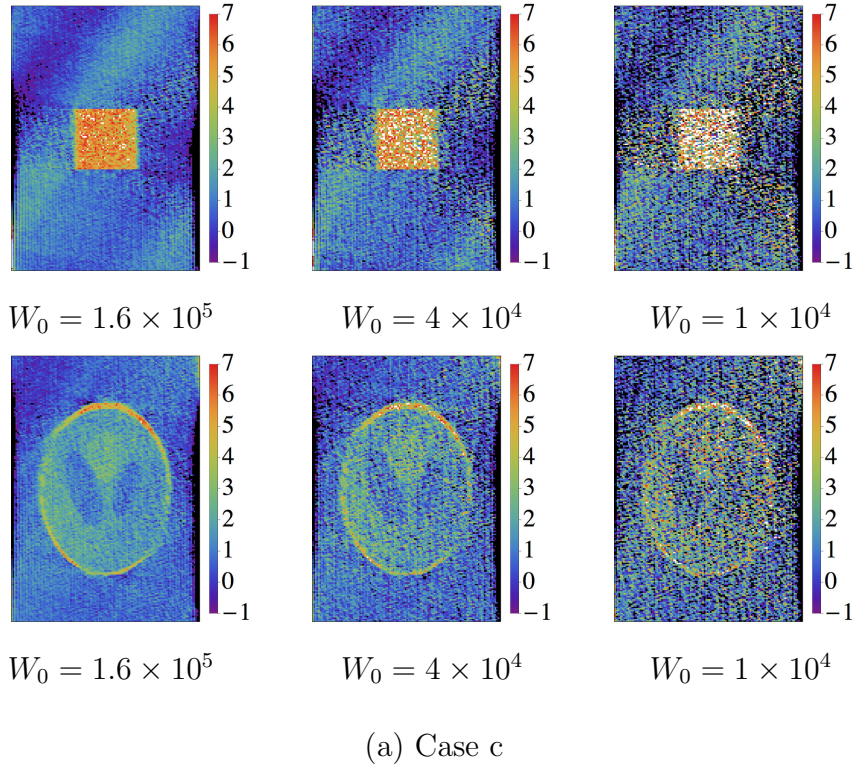$W_0 = 1.6 \times 10^5$ $W_0 = 4 \times 10^4$ $W_0 = 1 \times 10^4$

(b) Case d

Figure 7.6: Reconstruction with noise level $W_0 = 1.6 \times 10^5$, 4×10^4, 1×10^4 and $\lambda = 10^{-3}$ in experiment 1.
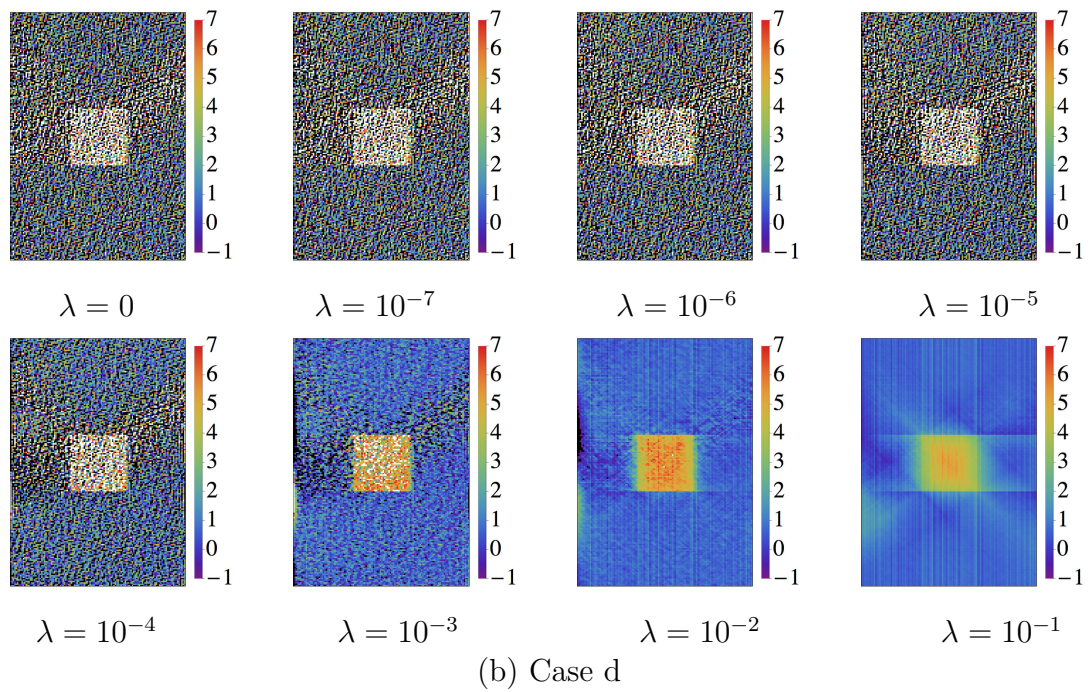
$\lambda = 0$ $\qquad$ $\lambda = 10^{-7}$ $\qquad$ $\lambda = 10^{-6}$ $\qquad$ $\lambda = 10^{-5}$

$\lambda = 10^{-4}$ $\qquad$ $\lambda = 10^{-3}$ $\qquad$ $\lambda = 10^{-2}$ $\qquad$ $\lambda = 10^{-1}$

(a) Case c

$\lambda = 0$ $\qquad$ $\lambda = 10^{-7}$ $\qquad$ $\lambda = 10^{-6}$ $\qquad$ $\lambda = 10^{-5}$

$\lambda = 10^{-4}$ $\qquad$ $\lambda = 10^{-3}$ $\qquad$ $\lambda = 10^{-2}$ $\qquad$ $\lambda = 10^{-1}$
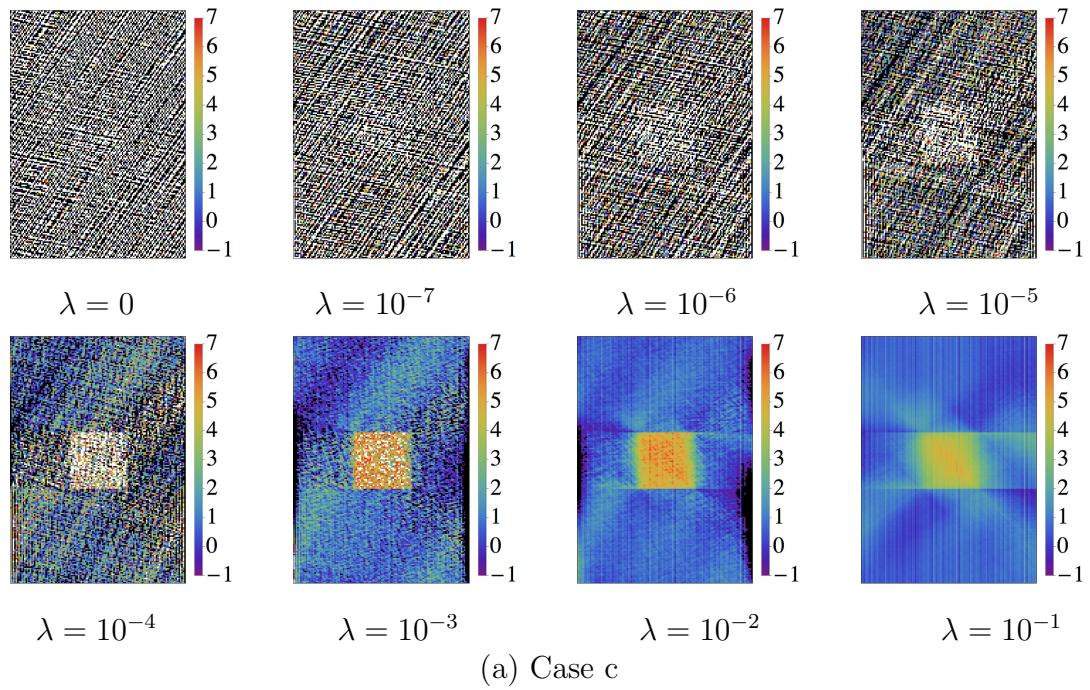
(b) Case d

Figure 7.7: Reconstruction with noise $W_0 = 4 \times 10^4$ and different regularization parameters.

reconstruction, though it does not meet the conditions in Section 4.2.

When we look at the results with noise added but no regularization, we can only see artifacts in all results except case (d). This validates that the conditions in Section 4.2 are essential for stable reconstruction. For reconstructions with $K = 2$, the artifacts have clear characteristic that their directions develop along the direction of $s_1 \hat{\mathbf{u}}_1 + s_2 \hat{\mathbf{u}}_2$, in the other word, the angle of artifacts is $\theta = (\theta_1 + \theta_2)/2$, which is the zero of $f(\theta)$ in Section 4.2.

We next turn to the results with regularization. Figure 7.6 shows the reconstruction for case (c) and (d) with different noise levels. The regularization parameter is $\lambda = 10^{-3}$. Regularization improves the results for both cases. The results in case (c) show bright and dark strips (in the left-top part of the results). but we don't have those artifacts in the results in (d).

The results turn worse when W_0 goes smaller. For $W_0 = 1 \times 10^4$, we could hardly see the details of the phantoms in Figure 7.6. Larger regularization can help remove the noise. Figure 7.7 shows the effect of regularization for different levels. For comparison, we show results for both case (c) and (d). For case (d), when λ is small, the reconstructions are almost the same. In Contrast, in case (c) the image quality continues getting worse as λ goes small, even for very small λ. When λ is comparably large, the results in two cases are very similar. The results get smooth as λ goes large. This is the reason that we'd like to involve other approaches to improve the quality of reconstruction while preserving the shape.
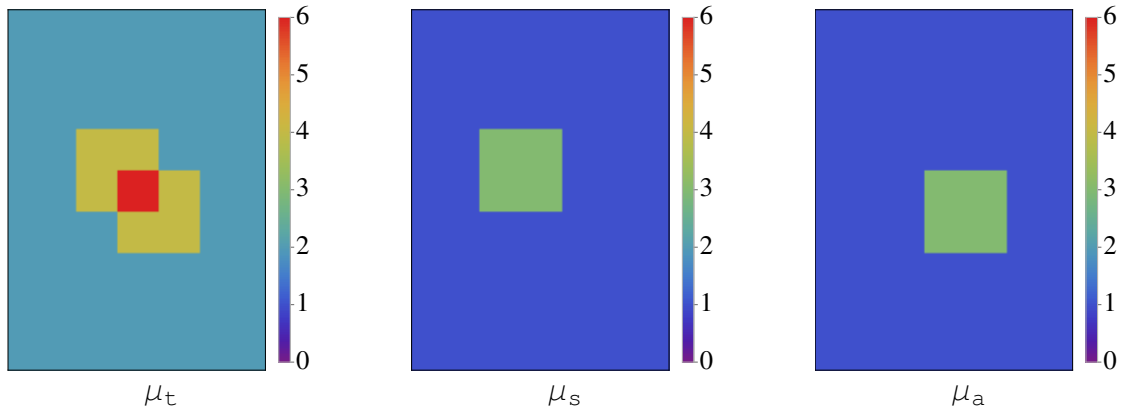
Figure 7.8: Phantoms 1: the total attenuation coefficient is two partly overlapped squares; the scattering and absorption coefficients are both squares in different locations. The background $\bar{\mu} = 0.01h^{-1}$ and $\bar{\mu}_s = 0.005h^{-1}$.

7.2.2 Reconstruction of scattering and absorption coefficients

We next turn to simultaneous reconstruction of scattering and absorption coefficients. We only perform experiments for $K = 3$. We use the same s_k and $\hat{\mathbf{u}}_k$ listed in case (d) in Table 7.1. Sampling is done in the same way as in last section. The scattering and absorption coefficients vary from $0.005h^{-1}$ to $0.03h^{-1}$. The attenuation coefficient is from $0.01h^{-1}$ to $0.06h^{-1}$. Scattering coefficient is positive for the entire domain with $\bar{\mu}_s = 0.005h^{-1}$. We also use two phantoms: In the first one, scattering and absorption coefficients are two overlapped squares; The second one has Shepp-Logan phantom as attenuation coefficient, while absorption and scattering coefficients are parts of the Shepp-Logan phantom. The phantoms are shown in Figure 7.8 and Figure 7.9.

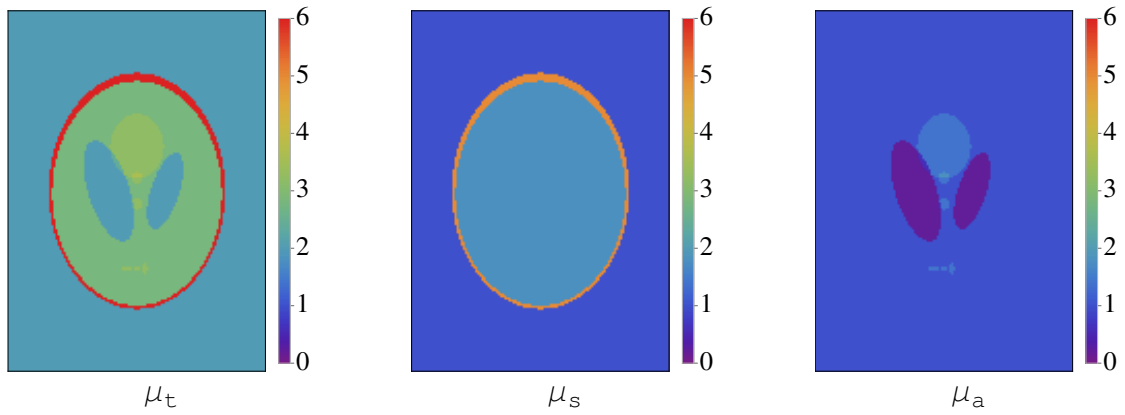Reconstruction without noise is shown in Figure 7.10. We have shown results for

Figure 7.9: Phantoms 2: the total attenuation coefficient is the Shepp-Logan phantom; the scattering and absorption coefficients are forming parts of the Shepp-Logan phantom. The background $\bar{\mu} = 0.01h^{-1}$ and $\bar{\mu}_s = 0.005h^{-1}$.
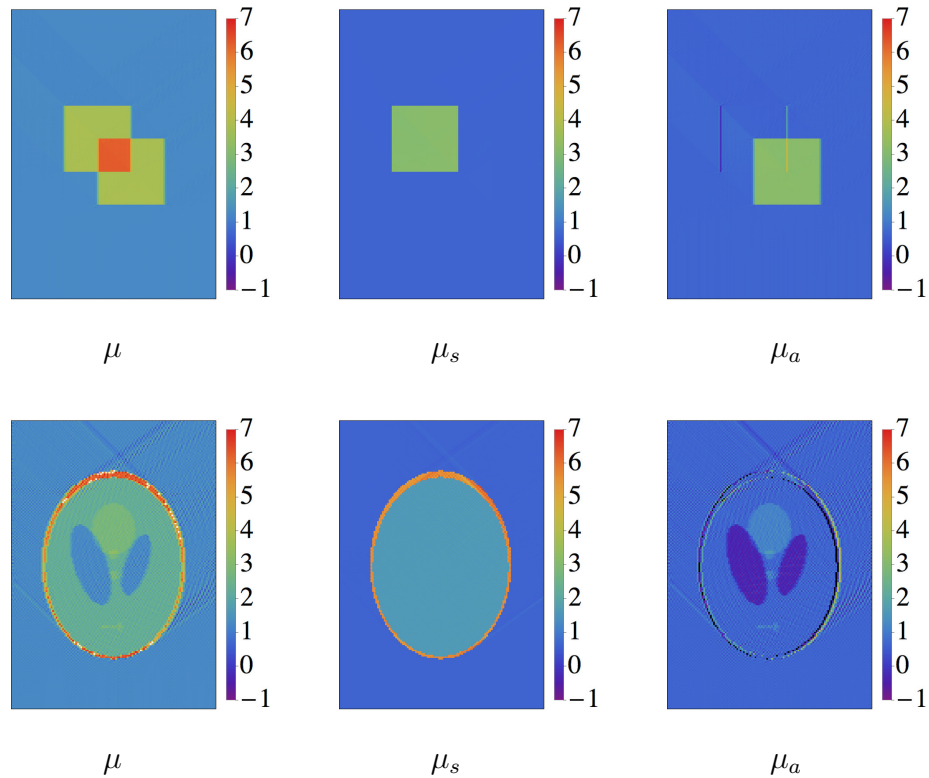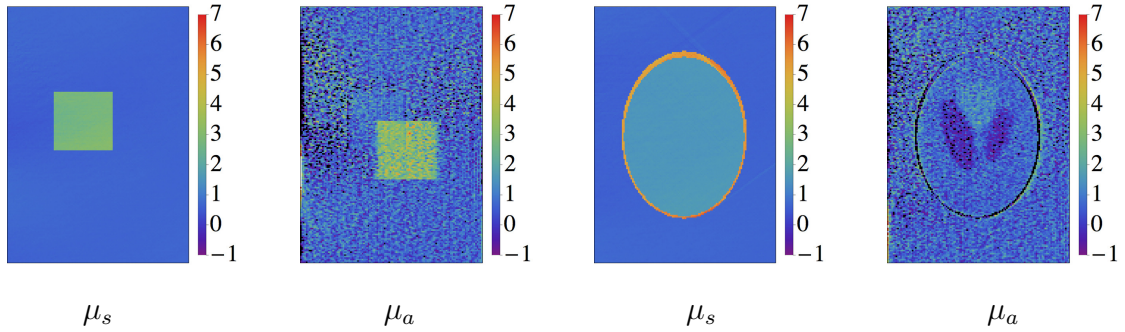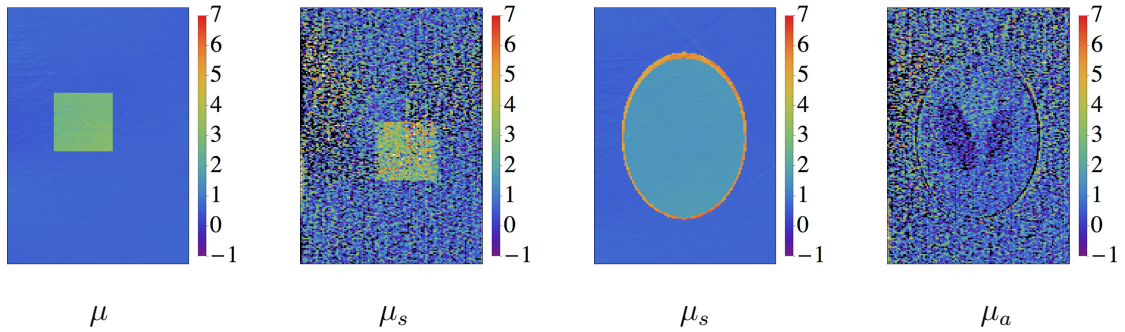

Figure 7.10: Reconstruction of μ_s and μ_a with no noise for both square and Shepp-Logan phantoms.

μ_s $\qquad$ μ_a $\qquad$ μ_s $\qquad$ μ_a

(a) $W_0 = 1.6 \times 10^5$

μ $\qquad$ μ_s $\qquad$ μ_s $\qquad$ μ_a

(b) $W_0 = 4 \times 10^4$

Figure 7.11: Reconstruction of μ_s and μ_a with noise for square and Shepp-Logan phantoms; (a) $W_0 = 1.6 \times 10^5$; (b) $W_0 = 4 \times 10^4$; Regularization parameter $\lambda = 10^{-3}$.

case (d) with both phantoms with different noise levels. All the results of μ_s have very clear shape. We can still see the artifacts and the shape of μ_s in the results of μ_a. When noise is added, as shown in Figure 7.11, reconstruction of μ_s are still very clear for even high noise level. However, reconstruction of μ_a are highly affected by noise. This is due to the way we generate our simulated data. In the 'inverse crime', μ_s is a weighted factor for Φ_{ij}, so it has a more significant effect to the simulated data.

7.2.3 Reconstruction with ballistic rays

For validating the approaches in improving image quality, experiments are done for only Case (d). The phantom used is the square shown in Figure 7.2 (a). The experiments are done with $R = 1$ and 2, where R is the number of ballistic rays. We use ballistic rays with directions $\hat{\mathbf{u}}_2$ and $\hat{\mathbf{u}}_3$ as listed in Table 7.1. The ballistic data is also computed by numerical integral, but no noise is added to it.

In Figure 7.12, we have reconstruction with noise level $W_0 = 4 \times 10^4$. In Figure 7.13, we have reconstruction with noise level $W_0 = 4 \times 10^4$ and regularization parameter $\lambda = 10^{-3}$. In the second row of each figure, we also show the difference between results in the first row. For both cases with $R = 1$ and 2, including ballistic data does not provide much improvement for reconstruction. The difference only appears near the boundary of the image. The differences between the results for $R = 2$ and the other two are more significant than the difference between no ballistic
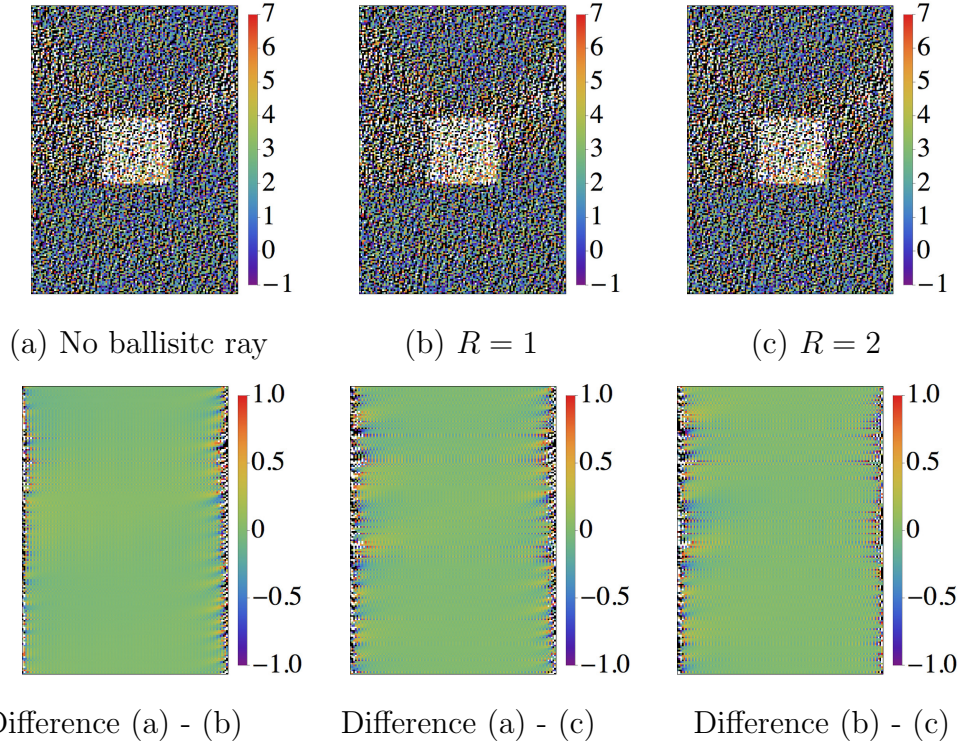
(a) No ballisitc ray   (b) $R = 1$ (c) $R = 2$

Difference (a) - (b)   Difference (a) - (c)   Difference (b) - (c)

Figure 7.12: Reconstruction with ballistic data for $W_0 = 4 \times 10^4$. The second row shows the differences between each images.

ray and $R = 1$. However, the improvement is still subtle.

7.2.4 Reconstruction with total variation regularization

For total variation regularized optimization, two issues need to be clarified. One is the initial point of iteration, and the other is the terminating criterion. We use the reconstruction in the first numerical experiment as the initial point. For termination, we require the difference of μ_j in two consecutive step to be small enough, i. e.

$$\|\mu_j - \mu_{j-1}\| \leq \delta \|\mu_0\| , \tag{7.2.1}$$

(a) No ballisitc ray   (b) $R = 1$ (c) $R = 2$

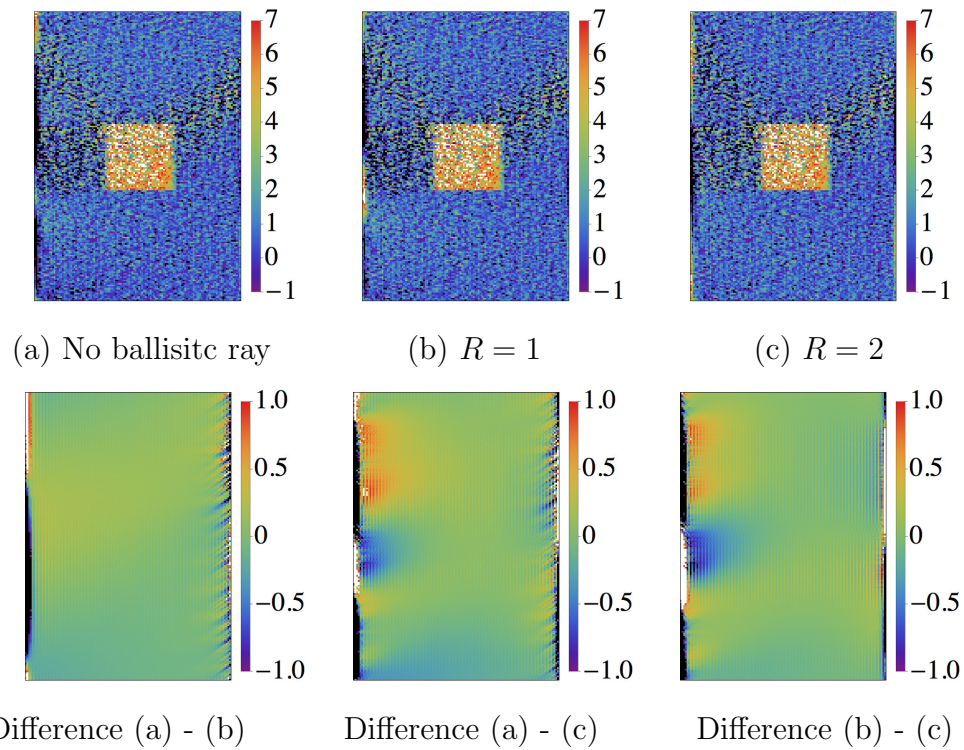Difference (a) - (b)   Difference (a) - (c)   Difference (b) - (c)

Figure 7.13: Reconstruction with ballistic data for $W_0 = 4 \times 10^4$ and regularization $\lambda = 10^{-3}$. The second row shows the differences between each images.

$$W_0 = 1.6 \times 10^5 \qquad W_0 = 4 \times 10^4 \qquad W_0 = 1 \times 10^4$$

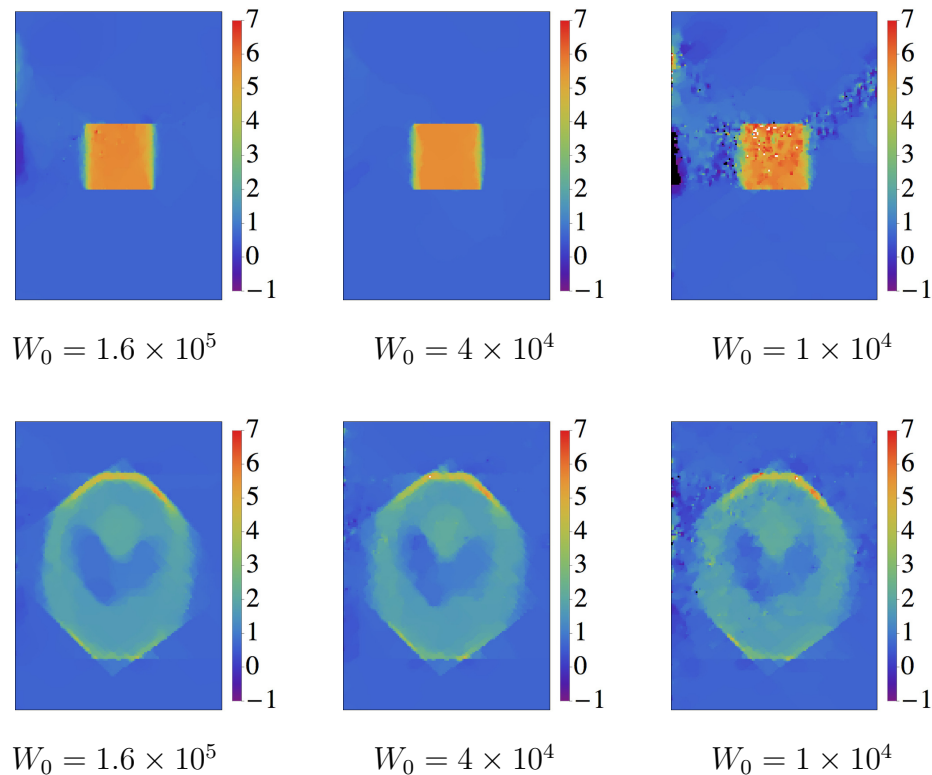$$W_0 = 1.6 \times 10^5 \qquad W_0 = 4 \times 10^4 \qquad W_0 = 1 \times 10^4$$

Figure 7.14: Reconstruction with total variation regularization for $W_0 = 1.6 \times 10^5$, 4×10^4 , 1×10^4 and regularization $\tau = 10^{-1}$.

where δ is some pre-chosen parameter. This criterion is used for both FISTA and the computation of P_L. We will present results for regularization parameter $\tau = 10^{-3}$ to 100.

Compared to the results with Tikhonov regularization, the reconstruction in Figure 7.14 indeed provides images with less noise. For the square phantom, we don't have so much artifacts as the results shown in Section 7.2.1, but we still have some artifacts when noise level is high (as shown in Figure 7.14 with $W_0 = 1 \times 10^4$). For Shepp-Logan phantom, the total variation regularization removes some details at the same time of removing noise. In Figure 7.15, we also present some results
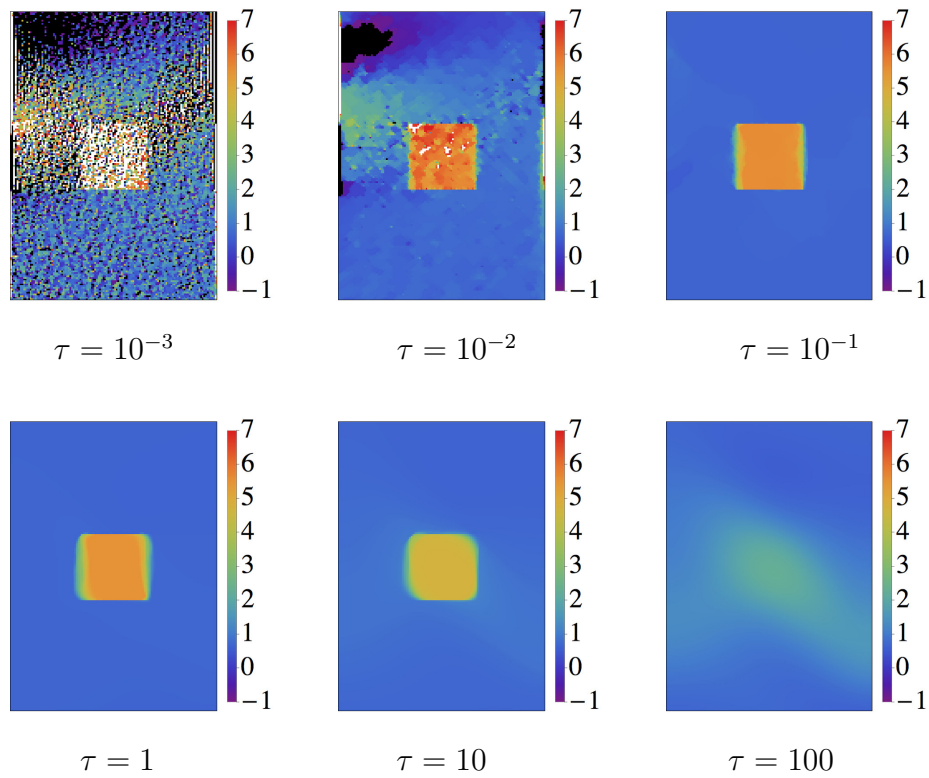
Figure 7.15: Reconstruction with total variation regularization for $W_0 = 4 \times 10^4$ and different regularization parameters.
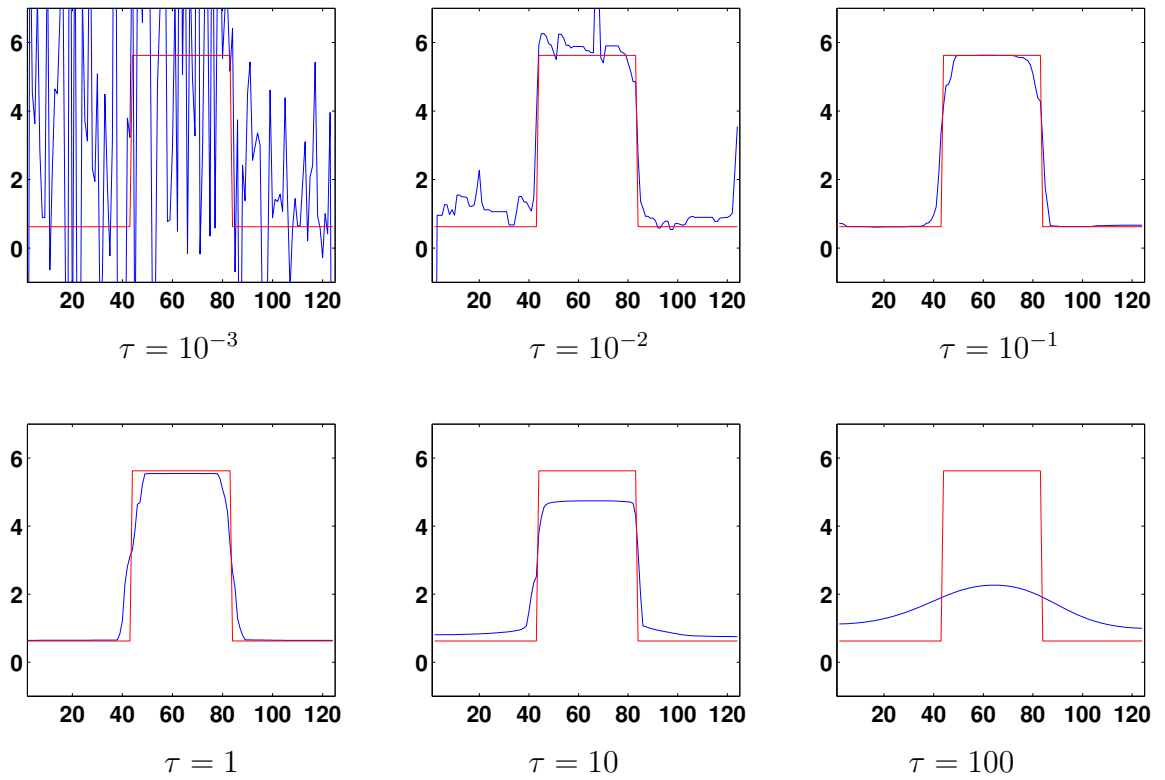
Figure 7.16: Cross sections of reconstruction with total variation regularization for $W_0 = 4 \times 10^4$ and different regularization parameters.

for $W_0 = 4 \times 10^4$ with different regularization parameters. When regularization parameter τ is between 10^{-1} to 10, the reconstruction have good shape as we expect.

7.2.5 Reconstruction with L^1 regularization

For L^1 regularization, we follow the same stopping criterion as described in the last section. The regularization parameter used here is chosen after comparing different choices. We will present results for regularization parameter $\tau = 10^{-4}$ to 10.

Regularization helps remove noise from the reconstruction. Even with higher noise level, the results are similar to the ones with lower noise level. As the same as in last section, results for square phantom are much better than the ones for Shepp-Logan phantom, because of the simple shape of square. However, compared to the results in experiment 1 and the ones with total variation regularization, the results shown in Figure 7.17 lose details near the boundary of the objects. Figure 7.18 shows that the results look like low-resolution images when regularization parameter is large. This is because the Haar wavelet is formed by a sequence of square-shaped functions.

$W_0 = 1.6 \times 10^5$ $W_0 = 4 \times 10^4$ $W_0 = 1 \times 10^4$

$W_0 = 1.6 \times 10^5$ $W_0 = 4 \times 10^4$ $W_0 = 1 \times 10^4$
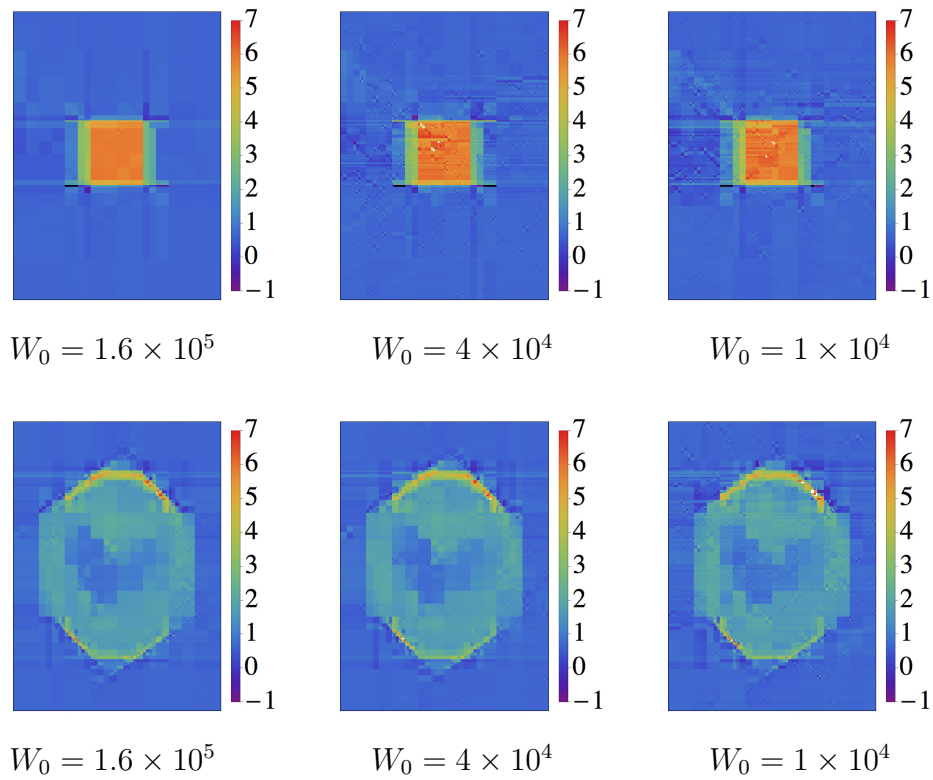
Figure 7.17: Reconstruction with L^1 regularization for $W_0 = 1.6 \times 10^5$, 4×10^4 , 1×10^4 and regularization $\tau = 10^{-1}$.

$\tau = 10^{-4}$ $\qquad$ $\tau = 10^{-3}$ $\qquad$ $\tau = 10^{-2}$

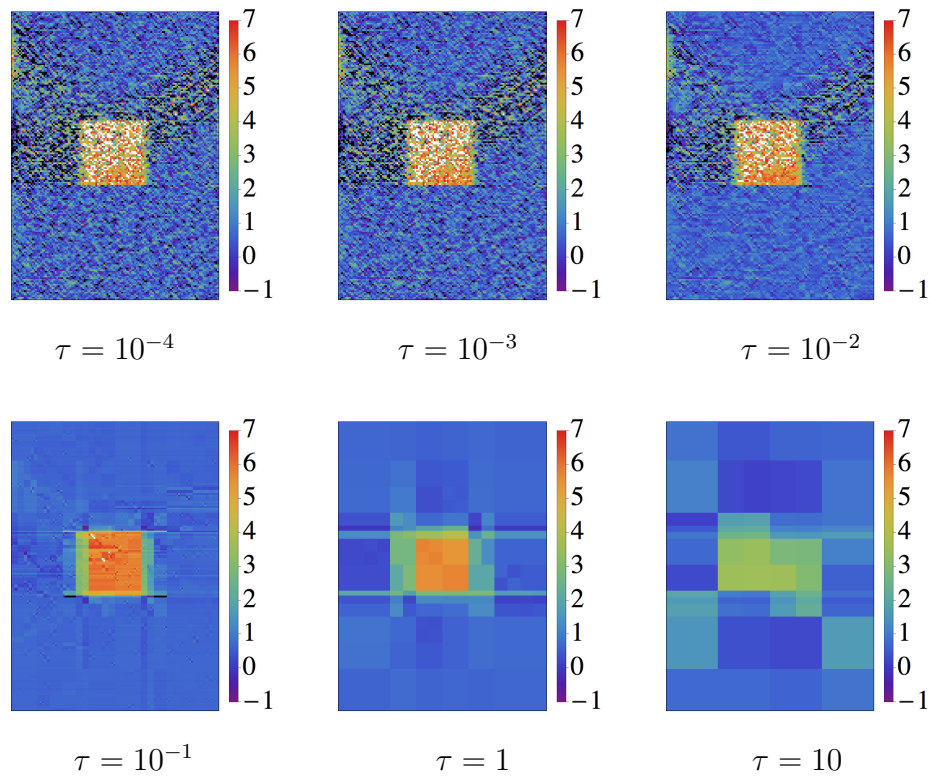$\tau = 10^{-1}$ $\qquad$ $\tau = 1$ $\qquad$ $\tau = 10$

Figure 7.18: Reconstruction with L^1 regularization for $W_0 = 4 \times 10^4$ and different regularization parameters.
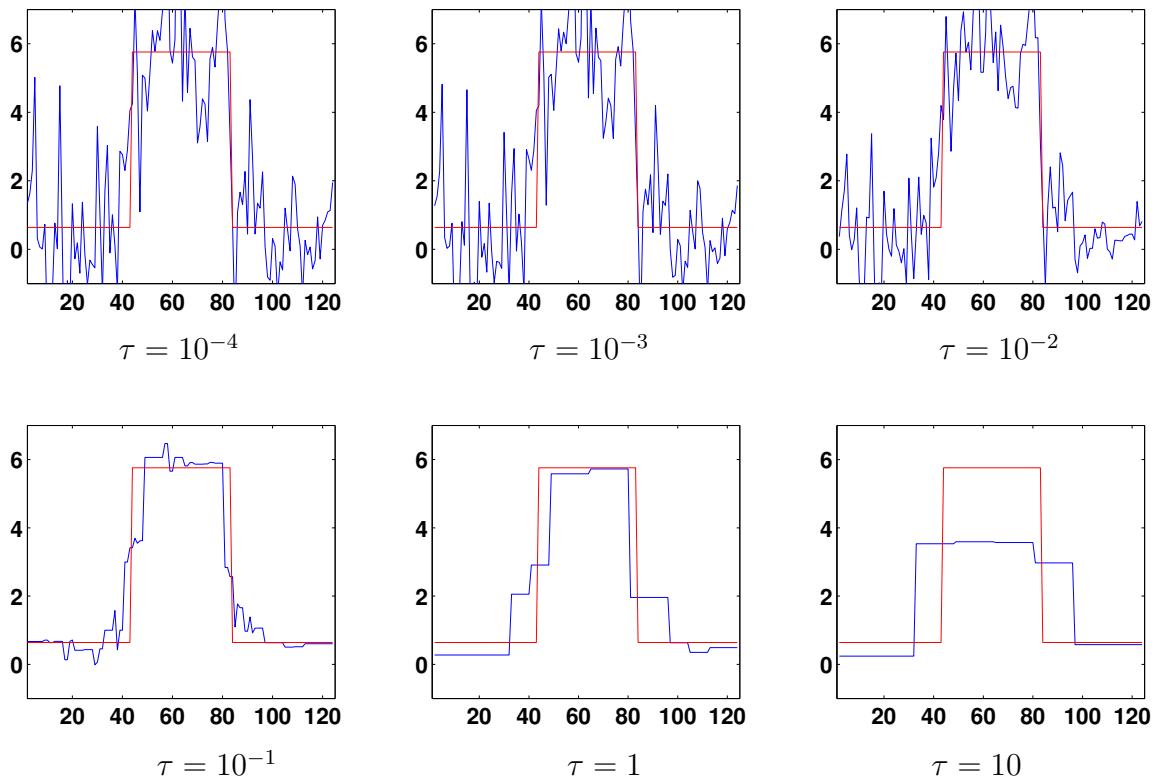
Figure 7.19: Cross sections of reconstruction with L^1 regularization for $W_0 = 4 \times 10^4$ and different regularization parameters.

Chapter 8

Discussion

In the chapters above, we have introduced star transform and its inverse integral problem, and talked about its potential application in finding attenuation coefficient in single scattering tomography. In Chapter 3, we presented a fast inversion algorithm for star transform in Fourier domain. The analysis of stability showed that inverse star transform is less ill-posed than inverse Broken ray Radon transform. We also provided necessary conditions for stable reconstruction in Chapter 4.

In later chapters, we provided an approach to recover scattering and absorption coefficients, which is theoretically feasible under our single scattering assumption. We also discussed approaches for removing noise from reconstructed attenuation coefficient: by including ballistic data for correction, utilizing total variation or L^1 regularization during reconstruction.

In the simulation chapter, we presented numerical experiments with simulated

data generated with 'inverse crime'. With different choices of parameters , we validated the necessary conditions mentioned in chapter 4. Experiments with noise also showed the necessity of regularization during reconstruction. The second experiment provided promising results of simultaneous reconstruction of scattering and absorption coefficients, even with high noise level.

Including ballistic data didn't provide the improvement that we expected. We can only use one or two sets of ballistic data, which may be much less than enough to have impact on previous reconstructions. However, this approach provides us a general idea how to combine scattered rays and ballistic rays together, while scattered rays are neglected in conventional CT and ballistic rays are not taken into account in SSOT. The hard constrained optimization problem gives us image that is consistent with the less noised ballistic rays. This scheme requires only small changes to the data collection and only a small amount of extra computation added to the inversion of star transform. Increasing the number of ballistic data could be a possible way to have greater improvement, but it requires significant modification to the system. In our geometry of single scattering tomography, we are not able to do that.

Total variation regularization helped us to obtain a sharp reconstruction for the square phantom in numerical experiment, but we still lost details for Shepp-Logan phantom. This suggested that further study is needed before it can be used in realistic application, because the internal structure of the object could be much

more complicated than simple objects like squares. The choice of regularization parameter needs to be studied in the future for optimal results. Compared to total variation regularization, L^1 regularization provided 'compressed image', for both square and Shepp-Logan phantom, when regularization parameter is large. This phenomenon is more obvious for the simpler square phantom than Shepp-Logan phantom. This is because we apply the L^1 regularization to the Haar coefficients of the image. Total variation regularization could be a better choice to be utilized in reconstruction.

More experiments for inverse star transform problem based on more realistic data is needed in the future in order to prevent the advantages we obtained from 'inverse crime'. Data generated from solving RTE or Monte Carlo could be better ones in the future research.

Bibliography

[1] Akysh, A. S. (2002). On a strong solution in the method of spherical harmonics for a nonstationary transport equation. Siberian Mathematical Journal, 43(4), 605-615.

[2] Beck, A., & Teboulle, M. (2009). A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM Journal on Imaging Sciences, 2(1), 183-202.

[3] Boas, D. (1997). A fundamental limitation of linearized algorithms for diffuse optical tomography. Optics Express, 1(13), 404-413.

[4] Boas, D. A., Brooks, D. H., Miller, E. L., DiMarzio, C. A., Kilmer, M., Gaudette, R. J., & Zhang, Q. (2001). Imaging the body with diffuse optical tomography. Signal Processing Magazine, IEEE, 18(6), 57-75.

[5] Boguna, M., Porra, J. M., & Masoliver, J. (2000). Continued fraction solution for the radiative transfer equation in three dimensions. Physical Review E, 61(6), 6248.

[6] Chambolle, A. (2004). An algorithm for total variation minimization and applications. Journal of Mathematical imaging and vision, 20(1-2), 89-97.

[7] Cormack, A. M. (1963). Representation of a function by its line integrals, with some radiological applications. Journal of applied physics, 34(9), 2722-2727.

[8] Cormack, A. M. (1981). The Radon transform on a family of curves in the plane. Proceedings of the American Mathematical Society, 83(2), 325-330.

[9] Culver, J. P., Choe, R., Holboke, M. J., Zubkov, L., Durduran, T., Slemp, A., ... & Yodh, A. G. (2003). Three-dimensional diffuse optical tomography in the parallel plane transmission geometry: evaluation of a hybrid frequency domain/continuous wave clinical system for breast imaging. Medical physics, 30(2), 235-247.

[10] Dorn, O. (1998). A transport-backtransport method for optical tomography. Inverse Problems, 14(5), 1107.

[11] Erdmann, R. C., & Siewert, C. E. (1968). Green's Functions for the One Speed Transport Equation in Spherical Geometry. Journal of Mathematical Physics, 9(1), 81-89.

[12] Florescu, L., Schotland, J. C., & Markel, V. A. (2009). Single-scattering optical tomography. Physical Review E, 79(3), 036607.

[13] Florescu, L., Markel, V. A., & Schotland, J. C. (2010). Single-scattering optical tomography: Simultaneous reconstruction of scattering and absorption. Physical Review E, 81(1), 016602.

[14] Florescu, L., Markel, V. A., & Schotland, J. C. (2011). Inversion formulas for the broken-ray Radon transform. Inverse Problems, 27(2), 025002.

[15] Frush, D. P. (2004, February). Review of radiation issues for computed tomography. In Seminars in Ultrasound, CT and MRI (Vol. 25, No. 1, pp. 17-24). WB Saunders.

[16] Gibson, A. P., Hebden, J. C., & Arridge, S. R. (2005). Recent advances in diffuse optical imaging. Physics in medicine and biology, 50(4), R1.

[17] Goldstein, T., ODonoghue, B., & Setzer, S. (2012). Fast alternating direction optimization methods. CAM report, 12-35.

[18] Hsieh, J. (2009, November). Computed tomography: principles, design, artifacts, and recent advances. Bellingham, WA: SPIE.

[19] Hull, E. L., & Foster, T. H. (2001). Steady-state reflectance spectroscopy in the P_3 approximation. JOSA A, 18(3), 584-599.

[20] Jiang, H., & Paulsen, K. D. (1995, May). Finite-element-based higher order diffusion approximation of light propagation in tissues. In Photonics West'95 (pp. 608-614). International Society for Optics and Photonics.

[21] Jiang, H. (1999). Optical image reconstruction based on the third-order diffusion equations. Optics Express, 4(8), 241-246.

[22] Katsevich, A., & Krylov, R. (2013). Broken ray transform: inversion and a range condition. Inverse Problems, 29(7), 075008.

[23] Kim, A. D. (2004). Transport theory for light propagation in biological tissue. JOSA A, 21(5), 820-827.

[24] Kuscer, I., & McCormick, N. J. (1991). Some analytical results for radiative transfer in thick atmospheres. Transport Theory and Statistical Physics, 20(5-6), 351-381.

[25] Maeland, E. (1998). Focusing aspects of the parabolic Radon transform. Geophysics, 63(5), 1708-1715.

[26] Morvidone, M., Nguyen, M. K., Truong, T. T., & Zaidi, H. (2010). On the V-line Radon transform and its imaging applications. Journal of Biomedical Imaging, 2010, 11.

[27] Nguyen, M. K., & Truong, T. T. (2010). Inversion of a new circular-arc Radon transform for Compton scattering tomography. Inverse Problems, 26(6), 065005.

[28] Norton, S. J. (1994). Compton scattering tomography. Journal of applied physics, 76(4), 2007-2015.

[29] Osher, S., Burger, M., Goldfarb, D., Xu, J., & Yin, W. (2005). An iterative regularization method for total variation-based image restoration. Multiscale Modeling & Simulation, 4(2), 460-489.

[30] van Rossum, M. V., & Nieuwenhuizen, T. M. (1999). Multiple scattering of classical waves: microscopy, mesoscopy, and diffusion. Reviews of Modern Physics, 71(1), 313.

[31] Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena, 60(1), 259-268.

[32] Siewert, C. E., & Maiorino, J. R. (1979). A point source in a finite sphere. Journal of Quantitative Spectroscopy and Radiative Transfer, 22(5), 435-439.

[33] Siewert, C. E. (1994). A radiative-transfer inverse-source problem for a sphere. Journal of Quantitative Spectroscopy and Radiative Transfer, 52(2), 157-160.

[34] Spott, T., & Svaasand, L. O. (2000). Collimated light sources in the diffusion approximation. Applied optics, 39(34), 6453-6465.

[35] Tarvainen, T., Vauhkonen, M., Kolehmainen, V., & Kaipio, J. P. (2006). Finite element model for the coupled radiative transfer equation and diffusion approximation. International Journal for Numerical Methods in Engineering, 65(3), 383-405.

[36] Tarvainen, T., Vauhkonen, M., Kolehmainen, V., & Kaipio, J. P. (2006). Finite element model for the coupled radiative transfer equation and diffusion approximation. International Journal for Numerical Methods in Engineering, 65(3), 383-405.

[37] Truong, T. T., & Nguyen, M. K. (2011). On new $\mathfrak{V}$-line Radon transforms in $\mathbb{R}^2$ and their inversion. Journal of Physics A: Mathematical and Theoretical, 44(7), 075206.

[38] Venugopalan, V., You, J. S., & Tromberg, B. J. (1998). Radiative transport in the diffusion approximation: an extension for highly absorbing media and small source-detector separations. Physical Review E, 58(2), 2395.