Summer 2013

Some Approximations to the Radiative Transfer Equation and Their Applications

Qiwei Sheng
University of Iowa

SOME APPROXIMATIONS TO THE RADIATIVE TRANSFER EQUATION AND THEIR APPLICATIONS

by

Qiwei Sheng

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy
degree in Applied Mathematical and Computational Sciences
in the Graduate College of
The University of Iowa

August 2013

Thesis Supervisor: Professor Weimin Han

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Qiwei Sheng

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Applied Mathematical and Computational Sciences at the August 2013 graduation.

Thesis Committee: _____
Weimin Han, Thesis Supervisor

_____
Bruce Ayati

_____
Laurent Jay

_____
Suely Oliveira

_____
David Stewart

To my parents, Laiping Sheng and Chengai Shi, who offered me unconditional love and support all the way since the beginning of my studies. This brief dedication is not enough room to express my gratitude that they are my parents.

ACKNOWLEDGEMENTS

I would like to thank several people for their help making this thesis possible.

First and foremost, my greatest thankfulness goes to my advisor, Professor Weimin Han, for his wise guidance, constant encouragement, and support throughout my time as a graduate researcher. His excellent mentorship will continue to shape my career long after I leave Iowa.

I want to personally thank Joe Eichholz for his help on the Matlab and C code used to generate the numerical examples presented in this thesis and valuable discussion on programming techniques.

I am grateful to Professor Kendall Atkinson, Professor Bruce Ayati, Professor Laurent Jay, Professor Suely Oliveira and Professor David Stewart for their time reviewing my work. They have positively influenced me in one way or another throughout my graduate studies.

Finally, I would like to thank my family for their endless encouragement. Their support has truly been a driving factor in my life.

TABLE OF CONTENTS

LIST OF TABLES v

LIST OF FIGURES vii

CHAPTER

1 INTRODUCTION 1

1.1 Motivation 1
1.2 Sobolev Spaces 4
1.3 The Radiative Transfer Equation 8

2 WELL-POSEDNESS OF THE FOKKER-PLANCK EQUATION IN A SCATTERING PROCESS 13

2.1 Introduction 13
2.2 Existence and Uniqueness 16
2.3 Maximum Principle and Positivity of Solution 22

3 DIFFERENTIAL APPROXIMATIONS AND THEIR APPLICATIONS TO INVERSE PROBLEMS 27

3.1 Differential Approximation of The Integral Operator 27
3.1.1 An Optimization Problem for Identifying the Parameters of Three Terms Approximation 38
3.2 Well Posedness of RT/DA Problems 40
3.3 An Iteration Method 45
3.4 Discretizations 47
3.5 Applications to Inverse Problems 48
3.6 Numerical Examples 54
3.6.1 Forward Problems 54
3.6.2 Forward Problems Using Optimized Parameters 64
3.6.3 Inverse Problems 65

4 CONCLUSIONS AND FURTHER WORK 76

REFERENCES 79

LIST OF TABLES

Table

3.1 L^2 error for several values of h with $n_\phi = 4$ and $n_\theta = 16$, linear basis . . . 59

3.2 L^2 error for several values of h with $n_\phi = 8$ and $n_\theta = 16$, linear basis . . . 59

3.3 L^2 error for several values of h with $n_\phi = 8$ and $n_\theta = 16$, quadratic basis 59

3.4 L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$ 60

3.5 L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$, quadratic basis 60

3.6 L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 12$, cubic basis . . . 61

3.7 L^2 error from different values of n_θ with $h = \frac{\sqrt{2}}{6}$, quadratic elements . . . 61

3.8 L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 12$, linear basis . . . 62

3.9 L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 12$, quadratic basis 63

3.10 L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$, linear basis . . . 63

3.11 L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$, quadratic basis 63

3.12 L^2 error for several values of h with $n_\phi = 4$ and $n_\theta = 16$, linear basis . . . 64

3.13 L^2 error for several values of h with $n_\phi = 4$ and $n_\theta = 20$, linear basis . . . 64

3.14 L^2 error for several values of h with $n_\phi = 8$ and $n_\theta = 16$, quadratic basis 65

3.15 L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 12$, quadratic basis 65

3.16 L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$, quadratic basis 66

3.17 Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, no noise 69

3.18 Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 5% noise 69

3.19 Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 10% noise 70

3.20 Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 30% noise 70

3.21 Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 50% noise 70

3.22 Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 70% noise 70

3.23 Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, no noise 71

3.24 Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 10% noise 71

3.25 Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 30% noise 72

3.26 Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, no noise 73

3.27 Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 20% noise 74

3.28 Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 40% noise 74

3.29 Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, no noise 74

3.30 True and reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 10% noise . 75

3.31 True and reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 20% noise . 75

LIST OF FIGURES

Figure

3.1 One-term approximation for $g = 0.9$ 35

3.2 Two-term approximation for $g = 0.9$ 36

3.3 Three-term approximation for $g = 0.9$ 37

3.4 The values of object function in terms of α_2 and α_3 (denoted by $a2$ and $a3$ in the Figure) for $g = 0.95$ 39

3.5 Angular mesh with 66 nodes. 55

3.6 Angular mesh with 66 nodes. 56

3.7 Domains R_1 and R_2 in different colors and pencil beams. 67

3.8 Domains R_1 and R_2 in different colors and pencil beams. 68

3.9 Domains R_1, R_2 and R_3 in different colors and pencil beams. 73

CHAPTER 1

INTRODUCTION

1.1 Motivation

Radiative transfer theory describes the interaction of radiation with scattering and absorbing media, which has wide applications in such areas as neutron transport, heat transfer, stellar atmospheres, optical molecular imaging, infrared and visible light in space and the atmosphere and so on. We refer to [1, 7, 19, 30, 43, 44] and references therein for details about this subject.

The main focus of our research is the application of the radiative transfer equation in biomedical imaging. One of potential applications is in novel techniques for the detection of breast cancer. X-ray mammography is currently the most prevalent imaging modality for screening and diagnosis of breast cancers. The use of mammography results in a 25–30% decreased mortality rate in screened women [33]. X-Ray Computerized Tomography offer good resolution but the parameters they measure do not vary enough from healthy to unhealthy tissues for proper discrimination. As a consequence, X-Ray CT sometime misses cancer tissues while sometimes gives false negative ([32, 38]). Some false negative and false positive diagnoses often led to missed cancers and inappropriate biopsies. The key factor that limits the success rate is the poor contrast between healthy and diseased tissues in the mammogram. Although X-ray CT of the breast can potentially improve diagnostic accuracy over mammography ([13, 14]), the state-of-the-art breast CT scanner is still based on the attenuation mechanism. As a result, the use of breast CT requires an intravenous

contrast medium and a high radiation dose, since elemental composition is almost uniform with little density variation in breast tissues. Still, it is rather difficult for breast CT to discern early-stage breast cancers.

Since the physical properties, especially the scattering capability, vary dramatically between tumor and healthy tissues, there is a significant difference of X-ray scattering behaviors between tumor and normal tissues ([42]). X-ray scattering is predominantly Rayleigh scattering ([29]). Every tissue component gives rise to a characteristic X-ray scattering pattern. Such scattering signals reveal critical information on ultra-fine features of cellular and sub-cellular structures, and bear unique diagnostic values of molecules, cells and their clusters which are also known as basic functional units of 100–200m in diameter ([6, 31]). Most importantly, X-ray scattering properties of tumors are significantly different from that of healthy tissue. Hence, X-ray Rayleigh scattering imaging provides a new contrast mechanism and would well complement attenuation-based X-ray imaging ([31, 9, 36, 37]). Thus, imaging techniques incorporating scattering information can potentially provide significantly higher contrast than conventional imaging modalities.

In principal, given a model of light propagation dependent upon scattering, we may recover the scattering by matching measurements of light leaving the domain in experiment to predictions from the model with varying scattering parameters. In stationary one-velocity case, the photon (neutron) intensity, a function of three space variables and two angular variables, is governed mathematically by the radiative transfer equation (RTE) (cf. [1],[28]). Since the use of the RTE in molecular imaging

will amount to solving an inverse problem with the RTE as the forward problem, fast and efficient methods of solving the RTE must be explored.

The RTE may be viewed as a hyperbolic type integro-differential equation. It is a serious challenge to solve the RTE problem numerically for a couple of reasons. First, it is a high dimensional problem. The spatial domain is three dimensional and the region for the angular variable is two dimensional. Second, when the RTE is discretized by the popular discrete-ordinate method, the integral term of the equation is approximated by a summation that involves all the numerical integration points on the unit sphere. Consequently, for the resulting discrete system, the desired locality property is not valid, and many of the solution techniques for solving sparse systems from discretization of partial differential equations cannot be applied efficiently to solve the discrete systems of RTE. Moreover, in applications involving highly forward-peaked media, which is typical in biomedical imaging, the phase function tends to be numerically singular. In such applications, it is even more difficult to solve RTE since accurate numerical solutions require a high resolution of the direction variable, leading to prohibitively large amount of computations. For these reasons, various approximations of RTE have been proposed in the literature, e.g., the delta-Eddington approximation ([25]), the Fokker-Planck approximation ([39, 40]), the Boltzmann-Fokker-Planck approximation ([41, 8]), the generalized Fokker-Planck approximation ([27]), the Fokker-Planck-Eddington approximation and the generalized Fokker-Planck-Eddington approximation ([16]). In [21] a family of differential approximations of the RTE(RT/DA) was proposed.

Certain type of Fokker-Planck equations in a domain or a smooth Riemannian manifold in the Euclidean space was discussed in the literature ([45, 17, 34, 10]). However, the equation discussed in this thesis is different from those in that there are two groups of independent variables: $\boldsymbol{x}$ in a domain and $\boldsymbol{\omega}$ from the unit sphere. In the equation we are going to discuss, the differential operator Δ^* is with respect to the angular variable $\boldsymbol{\omega}$ whereas ∇ is with respect to the spatial variable $\boldsymbol{x}$. For the equation in those papers, the differential operators Δ and ∇ are both with respect to the same independent variables. In Chapter 2, we study various properties of the Fokker-Planck equation such as well-posedness of the equation and positivity of the solution. In Chapter 3, we discuss a numerical method for solving GFP equations. The numerical method is based on a variational formulation involving even and odd components of the solution. We show the well-posedness of the variational formulation and develop a Galerkin method where spherical harmonics are used for the angular approximation and finite elements are used for spatial discretization. An iteration procedure is introduced to solve the problem and its convergence is shown. In Chapter 4, we study the RT/DA equations and investigate numerically the closeness of solutions of the RT/DA equations to that of the RTE. Finally we explore their applications to inverse problems.

1.2 Sobolev Spaces

In the study of partial differential equations and their numerical approximations, Sobolev spaces play a critical role. This section provides basic definitions and results concerning these spaces that will be used throughout the thesis.

Let X be a bounded domain in $\mathbb{R}^d$, and denote its boundary by ∂X. We will write a typical point in $\mathbb{R}^d$ as $\boldsymbol{x}$.

Given two arbitrary normed spaces A and B with norms $\|\cdot\|_A$ and $\|\cdot\|_B$ respectively. A function $f : A \to B$ is Lipschitz continuous if there exists a real constant C, such that, for all $x, y \in A$,

$$\|f(x) - f(y)\|_B \leq C\|x - y\|_A.$$

Any such C is referred to as a Lipschitz constant for the function f. The smallest constant is sometimes called the (best) Lipschitz constant. A domain X is said to be a Lipschitz domain if at each point $x \in X$ there exists $r > 0$ and a Lipschitz function $h : \mathbb{R}^{d-1} \to \mathbb{R}$ such that, upon rotating and relabeling the coordinate axes if necessary,

$$X \cap B(\boldsymbol{x}, r) = \{y \,|\, h(y_1, y_2, \cdots, y_{d-1}) < y_d)\} \cap B(\boldsymbol{x}, r).$$

In other words, near $\boldsymbol{x}$, ∂X is the graph of a Lipschitz function. In this thesis, $d = 3$.

For $p \in [1, \infty)$ we define $L^p(X)$ the space of all measurable functions u such that

$$\|u\|_{L^p(X)} = \left\{ \int_X |u|^p \, \mathrm{d}x \right\}^{1/p} < \infty.$$

For any $p \in [1, \infty)$, $L^p(X)$ is a Banach space, and $L^2(X)$ is a Hilbert space with inner product

$$(u, v)_{L^2(X)} = \int_X uv \, \mathrm{d}x.$$

We denote by $L^\infty(X)$ the space of all essentially bounded measurable functions and

use the norm

$$\|u\|_{L^\infty(X)} = \operatorname{ess\,sup}_{\boldsymbol{x}\in X} |u(\boldsymbol{x})|.$$

We denote the space of locally p-integrable functions as $L^p_{loc}(X)$. A function u is locally p-integrable if $v \in L^p(X')$ for any proper subset X' of X. In the case $p = 1$, we say that u is locally integrable.

Let $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$ be an n-tuple multiindex of order $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and $\alpha_i \geq 0$ for $1 \leq i \leq n$. If u is an m-times differentiable function then for any $|\alpha| \leq m$ we write

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}.$$

We refer to this function as the α^{th} partial derivative of u.

Define $C^m(\overline{X})$ as the space of all functions which are continuous on $\overline{X}$ and such that all of their partial derivatives of order at most m are also continuous on $\overline{X}$. In the special case $m = 0$ we write $C(\overline{X})$. If we endow $C^m(\overline{X})$ with the norm

$$\|u\|_{C^m(\overline{X})} = \max_{|\alpha|\leq m} \max_{\boldsymbol{x}\in\overline{X}} |D^\alpha u(\boldsymbol{x})|,$$

then $C^m(\overline{X})$ is Banach space. We also define

$$C^\infty(\overline{X}) = \cap_{m=0}^\infty C^m(\overline{X}).$$

Finally, we define $C_0^\infty(\overline{X})$ as the space of all compactly supported infinitely differentiable functions.

Suppose $u, v \in L^1_{loc}(X)$. Then v is called an α^{th}-weak partial derivative of u provided

$$\int_X u D^\alpha \phi \, \mathrm{d}x = (-1)^{|\alpha|} \int_X v\phi \, \mathrm{d}x, \qquad \forall \phi \in C_0^\infty(X),$$

in which case, we write

$$D^\alpha u = v.$$

A weak derivative, if it exists, is uniquely defined up to a set of measure zero with respect to the usual Lebesgue measure.

Now we are ready to define Sobolev space. For $1 \leq p \leq \infty$ and a non-negative integer k, the Sobolev space $W^{k,p}(X)$ consists of all functions $u \in L^1_{loc}(X) : X \to \mathbb{R}$ such that for each multiindex $\alpha \leq k$, $D^\alpha u$ exists in the weak sense and belongs to $L^p(X)$. We equip $W^{k,p}(X)$ with the norm

$$\|u\|_{W^{k,p}(X)} := \begin{cases} \left(\sum_{|\alpha|\leq k} \int_X |D^\alpha u|^p \, \mathrm{d}x\right)^{1/p} & p \in [1, \infty), \\ \sum_{|\alpha|\leq k} \operatorname{ess\,sup}_X |D^\alpha u| & p = \infty. \end{cases}$$

Then the Sobolev space $W^{k,p}(X)$ is a Banach space. In the case $p = 2$, it is customary to write

$$H^k(X) = W^{k,2}(X).$$

In this case, $H^k_0(X)$ is a Hilbert space with inner product

$$(u, v)_{H^k_0(X)} = \sum_{|\alpha|\leq k} \int_X D^\alpha u D^\alpha v \, \mathrm{d}x.$$

A function $u \in W^{k,p}(X)$ is only defined a.e. in X. Since ∂X has n-dimensional Lebesgue measure zero, there is no direct meaning we can give to the boundary value of u. We can resolve this problem by introducing a trace operator. The following theorem can be found in the literature, e.g. [12].

Theorem 1.2.1. *There exists a bounded linear operator*

$$T : W^{1,p}(X) \to L^p(\partial X)$$

such that

i. $Tu = u|_{\partial X}$ *if* $u \in W^{1,p}(X) \cap C(\overline{X})$ *and*

ii. $\|Tu\|_{L^p(\partial X)} \leq C\|u\|_{W^{k,p}(X)}$ *for some constant* $C > 0$.

We call T the trace operator and refer Tu to the generalized boundary value of u. We use u for both $u \in W^{1,p}(X)$ and its trace $Tu \in L^p(X)$. Its actual meaning is always clear from the context.

1.3 The Radiative Transfer Equation

Throughout this thesis, let X be a domain of the spatial variable, $X \subset \mathbb{R}^3$. Denote the boundary of X as ∂X and assume ∂X is Lipschitz. Note that this assumption implies that the unit outward normal $\boldsymbol{n}(x)$ exists a.e. on ∂X (cf. [3]). We will write Ω for $\mathbb{S}^2$, the boundary of the unit ball in $\mathbb{R}^3$. For each fixed direction $\boldsymbol{\omega} \in \Omega$, introduce a new Cartesian coordinate system (z_1, z_2, s) by the relations

$$\boldsymbol{x} = \boldsymbol{z} + s\boldsymbol{\omega}, \quad \boldsymbol{z} = z_1\boldsymbol{\omega}_1 + z_2\boldsymbol{\omega}_2,$$

where $(\boldsymbol{\omega}, \boldsymbol{\omega}_1, \boldsymbol{\omega}_2)$ is an orthonormal basis of $\mathbb{R}^3$, $z_1, z_2, s \in \mathbb{R}$. With respect to this new coordinate system, we denote by $X_{\boldsymbol{\omega}}$ the projection of X on the plane $s = 0$ in $\mathbb{R}^3$, and by $X_{\boldsymbol{\omega},\boldsymbol{z}}$ ($\boldsymbol{z} \in X_{\boldsymbol{\omega}}$) the intersection of the straight line $\{\boldsymbol{z} + s\boldsymbol{\omega}; s \in \mathbb{R}^3\}$ with X. We assume that the domain X is such that for any $(\boldsymbol{\omega}, \boldsymbol{z})$ with $\boldsymbol{z} \in X_{\boldsymbol{\omega}}$, $X_{\boldsymbol{\omega},\boldsymbol{z}}$ is the union of a nite number of line segments:

$$X_{\boldsymbol{\omega},\boldsymbol{z}} = \cup_{i=1}^{N(\boldsymbol{\omega},\boldsymbol{z})}\{\boldsymbol{z} + s\boldsymbol{\omega}; s \in (s_{i,-}, s_{i,+})\}.$$

Here $s_{i,\pm} = s_{i,\pm}(\boldsymbol{\omega}, \boldsymbol{z})$ depends on $\boldsymbol{\omega}$ and $\boldsymbol{z}$, and $\boldsymbol{x}_{i,\pm} := \boldsymbol{z} + s_{i,\pm}\boldsymbol{\omega}$ are the intersection points of the line $\{\boldsymbol{z} + s\boldsymbol{\omega}; s \in \mathbb{R}\}$ with ∂X. We further assume $\sup_{\boldsymbol{\omega},\boldsymbol{z}} N(\boldsymbol{\omega}, \boldsymbol{z}) < \infty$,

known as a generalized convexity condition. We then introduce the following subset of ∂X:

$$\partial X_{\boldsymbol{\omega},-} = \{\boldsymbol{z} + s_{i,-}\boldsymbol{\omega};\ 1 \le i \le N(\boldsymbol{\omega}, \boldsymbol{z}), \boldsymbol{z} \in X_{\boldsymbol{\omega}}\},$$

$$\partial X_{\boldsymbol{\omega},+} = \{\boldsymbol{z} + s_{i,+}\boldsymbol{\omega};\ 1 \le i \le N(\boldsymbol{\omega}, \boldsymbol{z}), \boldsymbol{z} \in X_{\boldsymbol{\omega}}\}.$$

It can be shown that if $\boldsymbol{n}(\boldsymbol{x}_{i,-})$ exists with $\boldsymbol{x}_{i,-} = \boldsymbol{z} + s_{i,-}\boldsymbol{\omega}$, then $\boldsymbol{n}(\boldsymbol{x}_{i,-}) \cdot \boldsymbol{\omega} \le 0$; if $\boldsymbol{x} \in \partial X$ and $\boldsymbol{n}(\boldsymbol{x}) \cdot \boldsymbol{\omega} \le 0$, then $\boldsymbol{x} \in \partial X_{\boldsymbol{\omega},-}$. Likewise, if $\boldsymbol{n}(\boldsymbol{x}_{i,+})$ exists with $\boldsymbol{x}_{i,+} = \boldsymbol{z} + s_{i,+}\boldsymbol{\omega}$, then $\boldsymbol{n}(\boldsymbol{x}_{i,+}) \cdot \boldsymbol{\omega} \ge 0$; if $\boldsymbol{x} \in \partial X$ and $\boldsymbol{n}(\boldsymbol{x}) \cdot \boldsymbol{\omega} \ge 0$, then $\boldsymbol{x} \in \partial X_{\boldsymbol{\omega},+}$. We can think $\partial X_{\boldsymbol{\omega},-}$ as the collection of all the points x in the boundary of X such that $\boldsymbol{\omega}$ is pointing towards the interior of X from the point $\boldsymbol{x}$. Similarly, $\partial X_{\boldsymbol{\omega},+}$ is the set of all points such that $\boldsymbol{\omega}$ is pointing away from X at the point x. Now define

$$\Gamma_- = \{(\boldsymbol{x}, \boldsymbol{\omega}) | \boldsymbol{x} \in \partial X_{\boldsymbol{\omega},-}, \boldsymbol{\omega} \in \Omega\}$$

and

$$\Gamma_+ = \{(\boldsymbol{x}, \boldsymbol{\omega}) | \boldsymbol{x} \in \partial X_{\boldsymbol{\omega},+}, \boldsymbol{\omega} \in \Omega\}$$

for the incoming and outgoing boundaries respectively. Both are subset of $\Gamma = \partial X \times \Omega$.

Denote by $\mathrm{d}\sigma(\boldsymbol{\omega})$ the innitesimal area element on the unit sphere Ω. If we introduce the spherical coordinate system by

$$\boldsymbol{\omega} = (\sin\theta\cos\psi, \sin\theta\sin\psi, \cos\theta)^T,\ 0 \le \theta \le \pi, 0 \le \psi \le 2\pi,$$

then $\mathrm{d}\sigma(\boldsymbol{\omega}) = \sin\theta\,\mathrm{d}\theta\,\mathrm{d}\psi$ and

$$\int_\Omega \varphi\,\mathrm{d}\sigma(\boldsymbol{\omega}) = \int_0^{2\pi} \mathrm{d}\psi \int_0^\pi \varphi\sin\theta\,\mathrm{d}\theta$$

for any function φ. We will need an integral operator S defined by

$$(Su)(\boldsymbol{x}, \boldsymbol{\omega}) = \int_\Omega k(\boldsymbol{x}, \boldsymbol{\omega} \cdot \boldsymbol{\omega}') u(\boldsymbol{x}, \boldsymbol{\omega}') \, \mathrm{d}\sigma(\boldsymbol{\omega}')$$

with k a nonnegative phase function satisfying the normalized condition:

$$\int_\Omega k(\boldsymbol{x}, \boldsymbol{\omega} \cdot \boldsymbol{\omega}') \, \mathrm{d}\sigma(\boldsymbol{\omega}') = 1 \qquad \forall \boldsymbol{x} \in X, \boldsymbol{\omega} \in \Omega.$$

In the applications of our interest the function k will be independent of $\boldsymbol{x}$. One well-known example is the Henyey-Greenstein phase function (cf. [24])

$$k(t) = \frac{1 - g^2}{4\pi(1 + g^2 - 2gt)^{3/2}}, \; t \in [-1, 1],$$

where the parameter $g \in (-1, 1)$. This particular phase function is popular in modeling scattering of light in a biological tissue.

We introduce a space of functions on $X \times \Omega$ analogous to the Sobolev spaces defined previously. Let $C^{(1,0)}(X \times \Omega)$ be the space of functions that are continuously differentiable in the spatial variable x for every fixed $\boldsymbol{\omega}$ and such that the directional derivative in the direction $\boldsymbol{\omega}$ is continuous with respect to all its arguments. Let ϕ be an arbitrary function in $C^{(1,0)}(X \times \Omega)$. For a function $u \in L^2(X \times \Omega)$ we define the weak directional derivative in direction $\boldsymbol{\omega}$, v, such that for all $\phi \in C^{(1,0)}(X \times \Omega)$ with compact support in X:

$$\int_{X \times \Omega} v\phi \, \mathrm{d}x \, \mathrm{d}\sigma(\boldsymbol{\omega}) = -\int_{X \times \Omega} u \, \boldsymbol{\omega} \cdot \nabla\phi \, \mathrm{d}x \, \mathrm{d}\sigma(\boldsymbol{\omega}).$$

If such a v exists, we will denote it as $\boldsymbol{\omega} \cdot \nabla u$. We define the space $H_2^1(X \times \Omega)$ as:

$$H_2^1(X \times \Omega) := \{v | v \in L^2(X \times \Omega) \text{ and } \boldsymbol{\omega} \cdot \nabla v \in L^2(X \times \Omega)\}.$$

If $u \in H_2^1(X \times \Omega)$ define its norm to be

$$\|u\|_{H_2^1(X\times\Omega)} := \|u\|_{L^2(X\times\Omega)} + \|\boldsymbol{\omega} \cdot \nabla u\|_{L^2(X\times\Omega)}.$$

We have the following useful property of the space $H_2^1(X \times \Omega)$ (cf. [1]):

Lemma 1.3.1. *$H_2^1(X \times \Omega)$ is complete, and $C^{1,0}(X \times \Omega)$ is dense in it.*

We also need function spaces $L^2(\Gamma_\pm)$ on $\Gamma_\pm$. They are defined to be spaces of functions u on $\Gamma_\pm$ such that

$$\|u\|_{L^2(\Gamma_\pm)} = \left[\int_\Omega \int_{X_{\boldsymbol{\omega},\pm}} \sum_{i=1}^{N(\boldsymbol{\omega},\boldsymbol{x})} |u(\boldsymbol{z} + s_{i,\pm}\boldsymbol{\omega}, \boldsymbol{\omega})|^2 \, \mathrm{d}z \, \mathrm{d}\sigma(\boldsymbol{\omega})\right]^{1/2} < \infty.$$

We have the following statement for the trace of an $H_2^1(X \times \Omega)$ function (cf. [1]). If $u \in H_2^1(X \times \Omega)$ and $u|_{\Gamma_-} \in L^2(\gamma_-)$, then $u|_{\Gamma_+} \in L^2(\gamma_+)$ and for some constant C depending only on X,

$$\|u\|_{L^2(\Gamma_+)} \le C[\|u\|_{H_2^1(X\times\Omega)} + \|u\|_{L^2(\Gamma_-)}]. \tag{1.3.1}$$

The radiative transfer equation considered is:

$$\boldsymbol{\omega} \cdot \nabla u(\boldsymbol{x}, \boldsymbol{\omega}) + \mu_t(\boldsymbol{x}) u(\boldsymbol{x}, \boldsymbol{\omega}) = \mu_s(\boldsymbol{x})(Su)(\boldsymbol{x}, \boldsymbol{\omega}) + f(\boldsymbol{x}, \boldsymbol{\omega}), \ (\boldsymbol{x}, \boldsymbol{\omega}) \in X \times \Omega, \tag{1.3.2}$$

$$u(\boldsymbol{x}, \boldsymbol{\omega}) = 0, \qquad (\boldsymbol{x}, \boldsymbol{\omega}) \in \Gamma_-. \tag{1.3.3}$$

Here $\mu_t = \mu_a + \mu_s$, μ_a is the macroscopic absorption cross section, μ_s is the macroscopic scattering cross section, and f is a source function. We assume these given functions satisfy

$$\mu_t, \mu_s \in L^\infty(X), \ \mu_s \ge 0 \text{ a.e. in } X, \tag{1.3.4}$$

$$\mu_t - \mu_s \geq c_0 \text{ in } X \text{ for some constant } c_0 > 0, \tag{1.3.5}$$

and

$$f \in L^2(X \times \Omega). \tag{1.3.6}$$

It is shown in [1] that the problem (1.3.2)–(1.3.3) has a unique solution $u \in H_2^1(X \times \Omega)$.

CHAPTER 2

WELL-POSEDNESS OF THE FOKKER-PLANCK EQUATION IN A SCATTERING PROCESS

2.1 Introduction

The numerical solution of the transport equation is challenging because of its high dimension and of its integro-differential form. In many applications, e.g., light propagation within biological tissues, there is a sharp peak in the forward scattering direction. Forward-peaked scattering corresponds to a sharp peak in the scattering phase function $k(\boldsymbol{\omega} \cdot \boldsymbol{\omega}')$ near $\boldsymbol{\omega} \cdot \boldsymbol{\omega}' = 1$. One well-known example is the Henyey-Greenstein phase function

$$k_{\mathrm{HG}}(\boldsymbol{\omega} \cdot \boldsymbol{\omega}') = \frac{1 - g^2}{4\pi(1 + g^2 - 2\, g\, \boldsymbol{\omega} \cdot \boldsymbol{\omega}')^{3/2}}, \tag{2.1.1}$$

where the parameter $g \in (-1, 1)$ is the anisotropy factor of the scattering medium which measures the strength of forward peakedness of the phase function. Typical values in biological tissues for g are in the range $0.9 \leq g \leq 0.99$, which correspond to quite highly forward-peaked scattering. This peak makes solving the transport equation even more difficult since the mesh size in such a calculation must be of the same magnitude as the mean free path, which, in this case, is very small. For this reason, there have been substantial efforts made to develop simpler approximations. The idea is to approximate the integral Boltzmann scattering operator

$$Su(\boldsymbol{x}, \boldsymbol{\omega}) := \int_{\Omega} k(\boldsymbol{\omega} \cdot \boldsymbol{\omega}') u(\boldsymbol{x}, \boldsymbol{\omega}')\, \mathrm{d}\boldsymbol{\omega}'.$$

One well-established example among these is the so-called Fokker-Planck approximation in which the scattering operator is approximated by a second-order differential operator, resulting in the Fokker-Planck equation:

$$-\frac{1}{2}\mu_{tr}(\boldsymbol{x})\Delta^* u(\boldsymbol{x},\boldsymbol{\omega}) + \boldsymbol{\omega}\cdot\nabla u(\boldsymbol{x},\boldsymbol{\omega}) + \mu_a(\boldsymbol{x})u(\boldsymbol{x},\boldsymbol{\omega}) = f(\boldsymbol{x},\boldsymbol{\omega}), \tag{2.1.2}$$

where Δ^* is the Laplace-Beltrami operator on Ω, $\mu_{tr}(\boldsymbol{x}) = (1-g)\,\mu_s(\boldsymbol{x})$ with

$$g = 2\pi\int_{-1}^{1} t\,\eta(t)\,\mathrm{d}t,$$

a measure of the degree of anisotropy. Note that the Henyey-Greenstein phase function k_{HG} is completely determined by the anisotropy parameter g, as is seen from the formula (2.1.1). Pomraning [39, 40] shows that the Fokker-Planck approximation is an asymptotic limit of the linear transport equation under certain conditions.

In the literature, one can find some papers ([45, 17, 34, 10]) that discuss properties of the following equation

$$-\Delta u + \nabla\cdot(u\,\boldsymbol{F}) = f \quad \text{in } X, \tag{2.1.3}$$

where X is either a domain or a smooth Riemannian manifold in the Euclidean space, $\boldsymbol{F}$ and f are given vector field and source function. This equation is also called a Fokker-Planck equation. However, the equation (2.1.2) discussed in this paper is different from (2.1.3) in that there are two groups of independent variables: $\boldsymbol{x}$ in a domain and $\boldsymbol{\omega}$ from the unit sphere. In (2.1.2), the differential operator Δ^* is with respect to the angular variable $\boldsymbol{\omega}$ whereas ∇ is with respect to the spatial variable $\boldsymbol{x}$. For the equation (2.1.3), the differential operators Δ and ∇ are both with respect to the same independent variables.

To derive of the Fokker-Planck equation, we first rewrite the radiative transfer equation as

$$\boldsymbol{\omega}\cdot\nabla u(x,\boldsymbol{\omega})+\mu_a u(x,\boldsymbol{\omega}) = \mu_s \int_\Omega k(\boldsymbol{\omega}\cdot\boldsymbol{\omega}')u(\boldsymbol{x},\boldsymbol{\omega}')\,\mathrm{d}\sigma(\boldsymbol{\omega}')-\mu_s u(x,\boldsymbol{\omega})+f(x,\boldsymbol{\omega}). \quad (2.1.4)$$

For k strongly peaked in the forward direction, the integral operator in (2.1.4) can be simplified by expanding $u(x,\boldsymbol{\omega})$ in a Taylor series with respect to $\boldsymbol{\omega}'$ about $\boldsymbol{\omega}'=\boldsymbol{\omega}$,

$$u(x,\boldsymbol{\omega}') = u(x,\boldsymbol{\omega}) + (\boldsymbol{\omega}'-\boldsymbol{\omega})\cdot\nabla^* u(x,\boldsymbol{\omega}) + \frac{1}{2}[(\boldsymbol{\omega}'-\boldsymbol{\omega})\cdot\nabla^*]^2 u(x,\boldsymbol{\omega}) + O\left[(\boldsymbol{\omega}'-\boldsymbol{\omega})^3\right]. \quad (2.1.5)$$

When (2.1.5) is used in the integral in (2.1.4), the first term cancels $\mu_s u$, the second term vanishes because it is odd, and the third term yields $\frac{1}{2}\mu_{tr}\Delta^* u$. Pomraning [39] and Larsen [26] provide complete derivations of this approximation.

We can also derive this result by approximating the sharply peaked scattering phase function as a delta function plus a higher-order correction:

$$k(\boldsymbol{\omega}\cdot\boldsymbol{\omega}') \approx a_0\delta(\boldsymbol{\omega}'-\boldsymbol{\omega}) + a_1\delta''(\boldsymbol{\omega}'-\boldsymbol{\omega}). \quad (2.1.6)$$

Here we interpret δ'' as the kernel corresponding to the spherical Laplace-Beltrami operator so that

$$\int_\Omega \delta''(\boldsymbol{\omega}'-\boldsymbol{\omega})u(\boldsymbol{x},\boldsymbol{\omega})\,\mathrm{d}\sigma(\boldsymbol{\omega}) = \Delta^* u(\boldsymbol{\omega}). \quad (2.1.7)$$

Terms involving derivatives of odd order must be zero to preserve rotational invariance of the scattering operator. The eigenvalues of the Laplace-Beltrami operator are $-n(n+1)$ and the associated eigenfunctions may be taken as spherical harmonics $Y_{n,m}, 0\le n\le\infty, -n\le m\le n$. The eigenvalues for integral operator S with k taken

to be the Henyey-Greenstein phase function are g^n and the associated eigenfunctions are also $Y_{n,m}$. We choose $a_0 = 1$ and $a_1 = (1-g)/2$ so that the resulting operator has the same first two eigenvalues as to the integral operator S. Substituting (2.1.6) into (2.1.4) using these coefficients we obtain (2.1.2) again.

2.2 Existence and Uniqueness

Despite the fact that quite a few papers discuss the Fokker-Planck equation, a rigorous study of its solution existence and uniqueness appears to be missing. The purpose of this section is to fill this gap.

We denote by Γ the boundary of $X \times \Omega$: $\Gamma := \partial(X \times \Omega) = \partial X \times \Omega$. In this chapter we assume the boundary ∂X of the domain X has the C^1 smoothness and denote by $\boldsymbol{\nu}(\boldsymbol{x})$ the unit outward normal vector at a point $\boldsymbol{x} \in \partial X$. In section 1.3, we give the definition of $\Gamma_\pm$ assuming the boundary ∂X is Lipschitz. When the boundary ∂X of X is C^1, we can simplify the definition of the incoming boundary Γ_- as

$$\Gamma_- = \{(\boldsymbol{x}, \boldsymbol{\omega}) \in \Gamma \mid \boldsymbol{\nu}(\boldsymbol{x}) \cdot \boldsymbol{\omega} < 0\},$$

which is equivalent to the definition given in Section 1.3 in case of C^1 boundary. In addition to the incoming boundary Γ_-, we further introduce the boundary subsets

$$\Gamma_+ = \{(\boldsymbol{x}, \boldsymbol{\omega}) \in \Gamma \mid \boldsymbol{\nu}(\boldsymbol{x}) \cdot \boldsymbol{\omega} > 0\}, \quad \Gamma_0 = \{(\boldsymbol{x}, \boldsymbol{\omega}) \in \Gamma \mid \boldsymbol{\nu}(\boldsymbol{x}) \cdot \boldsymbol{\omega} = 0\};$$

Γ_+ being known as the outgoing boundary. There holds the decomposition $\Gamma = \Gamma_0 \cup \Gamma_+ \cup \Gamma_-$.

We consider the BVP

$$Au = f \qquad \text{in } X \times \Omega, \tag{2.2.1a}$$

$$u = 0 \qquad \text{on } \Gamma_-. \tag{2.2.1b}$$

Here A is a partial differential operator of the form given by the left side of (2.1.2):

$$Au(\boldsymbol{x}, \boldsymbol{\omega}) = -a_1(\boldsymbol{x})\Delta^* u(\boldsymbol{x}, \boldsymbol{\omega}) + \boldsymbol{\omega} \cdot \nabla u(\boldsymbol{x}, \boldsymbol{\omega}) + a_2(\boldsymbol{x}) u(\boldsymbol{x}, \boldsymbol{\omega}), \tag{2.2.2}$$

where Δ^* is the Laplace-Beltrami operator on the unit sphere Ω. We assume

$$a_1, a_2 \in L^\infty(X), \quad a_1, a_2 \geq c_0 \text{ in } X \text{ for some constant } c_0 > 0, \tag{2.2.3}$$

and

$$f \in L^2(X \times \Omega).$$

For the Fokker-Planck equation (2.1.2), these assumptions are naturally valid; the second part of (2.2.3) is automatically satisfied as long as the scattering and absorption effects are not negligible.

We now introduce a weak formulation for a boundary value problem (2.2.1). For this purpose, let $u \in C^2(\overline{X} \times \Omega)$ be a function satisfying (2.2.1). Multiply the equation (2.2.1a) by a test function $v \in C^1(\overline{X} \times \Omega)$, integrate over $X \times \Omega$, and apply the boundary condition (2.2.1b) to find

$$a(u, v) = (f, v) \tag{2.2.4}$$

in which,

$$a(u, v) = \int_{X \times \Omega} (a_1 \nabla^* u \cdot \nabla^* v - u\, \boldsymbol{\omega} \cdot \nabla v + a_2 uv)\, \mathrm{d}x\, \mathrm{d}\sigma(\boldsymbol{\omega}) + \int_{\Gamma_+} \boldsymbol{\nu} \cdot \boldsymbol{\omega}\, uv\, \mathrm{d}S\, \mathrm{d}\sigma(\boldsymbol{\omega}),$$

$$(f, v) = \int_{X \times \Omega} fv\, \mathrm{d}x\, \mathrm{d}\sigma(\boldsymbol{\omega}),$$

where $\mathrm{d}S$ is the infinitesimal area element on ∂X, and $(\cdot,\cdot)$ denotes the inner product in $L^2(X\times\Omega)$. The bilinear form $a(\cdot,\cdot)$ defined on $C^1(\overline{X}\times\Omega)\times C^1(\overline{X}\times\Omega)$ is not symmetric. So we introduce another bilinear form

$$\overline{a}(v,w)=\int_{X\times\Omega}\left(a_1\nabla^* v\cdot\nabla^* w+a_2 vw\right)\,\mathrm{d}x\,\mathrm{d}\sigma(\boldsymbol{\omega})+\frac{1}{2}\int_{\Gamma}|\boldsymbol{\nu}\cdot\boldsymbol{\omega}|vw\,\mathrm{d}S\,\mathrm{d}\sigma(\boldsymbol{\omega}) \quad (2.2.5)$$

which symmetries $a(\cdot,\cdot)$:

$$a(v,v)=\overline{a}(v,v)\quad\forall\, v\in C^1(\overline{X}\times\Omega).$$

Under the condition (2.2.3), $\overline{a}(\cdot,\cdot)$ defines an inner product on $C^1(\overline{X}\times\Omega)$. We introduce the completion of the space $C^1(\overline{X}\times\Omega)$ with respect to the inner product $\overline{a}(\cdot,\cdot)$:

$$V_1:=\left\{v\in L^2(X\times\Omega)\mid|\nabla^* v|\in L^2(X\times\Omega),\int_{\Gamma}|\boldsymbol{\nu}\cdot\boldsymbol{\omega}|v^2\,\mathrm{d}S\,\mathrm{d}\sigma(\boldsymbol{\omega})<\infty\right\}, \quad (2.2.6)$$

with the norm

$$\|v\|_{V_1}=\sqrt{\overline{a}(v,v)},$$

which is equivalent to the standard norm

$$\left[\int_{X\times\Omega}\left(|\nabla^* v|^2+v^2\right)\,\mathrm{d}x\,\mathrm{d}\sigma(\boldsymbol{\omega})+\int_{\Gamma}|\boldsymbol{\nu}\cdot\boldsymbol{\omega}|v^2\,\mathrm{d}S\,\mathrm{d}\sigma(\boldsymbol{\omega})\right]^{1/2}.$$

We further introduce a subspace of V_1:

$$V_2:=\left\{v\in V_1\mid\boldsymbol{\omega}\cdot\nabla v\in L^2(X\times\Omega)\right\}, \quad (2.2.7)$$

with the norm

$$\|v\|_{V_2}=\left[\|v\|_{V_1}^2+\int_{X\times\Omega}|\boldsymbol{\omega}\cdot\nabla v|^2\,\mathrm{d}x\,\mathrm{d}\sigma(\boldsymbol{\omega})\right]^{1/2}.$$

Here, all the derivatives are understood to be the generalized (weak) derivatives.

We can extend $a(\cdot,\cdot)$ continuously with respect to its first argument in V_1 and its second argument in V_2. Denote the extension again by $a(\cdot,\cdot)$. Similarly, the bilinear form $\overline{a}(\cdot,\cdot)$ is extended continuously to $V_1 \times V_1$. Moreover,

$$a(v,v) = \overline{a}(v,v) \quad \forall\, v \in V_2. \tag{2.2.8}$$

With the above preparation, we can define a weak solution of the BVP (2.2.1) as follows.

Definition 2.2.1. We say that $u \in V_1$ is a weak solution of the BVP (2.2.1) if

$$a(u,v) = (f,v) \quad \forall\, v \in V_2.$$

We now study the weak formulation given in Definition 2.2.1. First observe that the bilinear form is bounded: for some appropriate constant C,

$$|a(u,v)| \le C\,\|u\|_{V_1}\|v\|_{V_2} \quad \forall\, u \in V_1,\, v \in V_2.$$

This is proved by applying the Cauchy-Schwarz inequality. In fact, we have

$$\begin{aligned}
|a(u,v)| \le & \int_{X\times\Omega} \left(a_1|\nabla^* u\cdot\nabla^* v| + |u\,\boldsymbol{\omega}\cdot\nabla v| + |a_2 uv|\right)\,\mathrm{d}x\,\mathrm{d}\sigma(\boldsymbol{\omega}) \\
& + \int_{\Gamma_+} |\boldsymbol{\nu}\cdot\boldsymbol{\omega}\,uv|\,\mathrm{d}S\,\mathrm{d}\sigma(\boldsymbol{\omega}) \\
\le & \int_{X\times\Omega} (a_1|\nabla^* u|^2\,\mathrm{d}x\,\mathrm{d}\omega \int_{X\times\Omega} a_1|\nabla^* v|^2\,\mathrm{d}x\,\mathrm{d}\sigma(\boldsymbol{\omega}) \\
& + \int_{X\times\Omega} u^2\,\mathrm{d}x\,\mathrm{d}\omega \int_{X\times\Omega} |\boldsymbol{\omega}\cdot\nabla v|^2\,\mathrm{d}x\,\mathrm{d}\sigma(\boldsymbol{\omega}) \\
& + \int_{X\times\Omega} a_2\,u^2\,\mathrm{d}x\,\mathrm{d}\sigma(\boldsymbol{\omega}) \int_{X\times\Omega} a_2\,v^2\,\mathrm{d}x\,\mathrm{d}\sigma(\boldsymbol{\omega}) \\
& + \int_{\Gamma} |\boldsymbol{\nu}\cdot\boldsymbol{\omega}|\,u^2\,\mathrm{d}S\,\mathrm{d}\sigma(\boldsymbol{\omega}) \int_{\Gamma} |\boldsymbol{\nu}\cdot\boldsymbol{\omega}|\,v^2\,\mathrm{d}S\,\mathrm{d}\sigma(\boldsymbol{\omega}) \\
\le & C\,\|u\|_{V_1}\|v\|_{V_2}.
\end{aligned}$$

Thus, for any fixed $v \in V_2$, $u \mapsto a(u,v)$ is a bounded linear functional on V_1. Therefore, by the Riesz representation theorem (e.g., [3, Subsection 2.5.2]), there exists a unique element Tv of V_1 such that

$$a(u,v) = \overline{a}(u,Tv) \quad \forall\, u \in V_1. \tag{2.2.9}$$

We claim $T : V_2 \to V_1$ is a bounded linear operator. Indeed if $\lambda_1, \lambda_2 \in \mathbb{R}$ and $v_1, v_2 \in V_2$, we see for each $u \in V_1$ that

$$\begin{aligned}\overline{a}(u, T(\lambda_1 v_1 + \lambda_2 v_2)) &= a(u, \lambda_1 v_1 + \lambda_2 v_2) \quad \text{(by (2.2.9))} \\ &= \lambda_1 a(u, v_1) + \lambda_2 a(u, v_2) \\ &= \overline{a}(u, \lambda_1 T v_1 + \lambda_2 T v_2).\end{aligned}$$

So T is linear. Furthermore

$$\|Tv\|_{V_1}^2 = \overline{a}(Tv, Tv) = a(Tv, v) \leqslant C\|Tv\|_{V_1}\|v\|_{V_2}.$$

Consequently $\|Tv\|_{V_1} \leqslant C\,\|v\|_{V_2}$ for all $v \in V_2$, and so $T : V_2 \to V_1$ is bounded.

Since $\overline{a}(v,Tv) = \overline{a}(v,v)$ by (2.2.8) and (2.2.9), we have

$$\|v\|_{V_1}^2 = \overline{a}(v,v) = \overline{a}(v,Tv) \leq \|v\|_{V_1}\|Tv\|_{V_1},$$

therefore

$$\|v\|_{V_1} \leq \|Tv\|_{V_1} \quad \text{if } v \in V_2. \tag{2.2.10}$$

So, T is a one-to-one mapping from V_2 into V_1.

Obviously, $\{Tv \mid v \in V_2\} \subset V_1$. However $\{Tv \mid v \in V_2\}$ is not a complete subset of V_1. Furthermore, T is injective from V_2 while not subjective to V_1. So we

are trying to enlarge V_2 and reduce V_1 to further make T a bijective map. To achieve this, we first take to completion of $\{Tv \mid v \in V_2\}$ and denote it by $V_{1,T}$. Then $V_{1,T}$ is a closed subset of V_1. Next we add some elements to V_2 and form a new set so that the image of this new set under map T is $V_{1,T}$. To do this, let $\{v_n\} \subset V_2$ be such that $\{Tv_n\}$ is a Cauchy sequence in V_1 with the limit $w = \lim_{n\to\infty} Tv_n \in V_{1,T}$. Then $\{v_n\}$ is a Cauchy sequence in V_1 by (2.2.10) and it has a limit $v \in V_1$. Notice that v may not be in V_2 and the value of Tv can be defined by $Tv := w = \lim_{n\to\infty} Tv_n$. With these in mind, we introduce subspaces

$$V_{1,T} := \text{completion of } \{Tv \mid v \in V_2\} \text{ in } V_1,$$

$$V_{2,T} := \text{completion of } V_2 \text{ with respect to } v \mapsto \|Tv\|_{V_1}.$$

The operator T is then extended from V_2 to $V_{2,T}$ by $Tv := \lim_{n\to\infty} Tv_n$. $V_{1,T}$ is a closed subspace of V_1, $V_2 \subset V_{2,T} \subset V_1$, and T can be extended to an isomorphism from $V_{2,T}$ onto $V_{1,T}$ (the extension is again denoted by T). Moreover, the equality

$$a(Tv, w) = \overline{a}(Tv, Tw)$$

is extended from $V_2 \times V_2$ to $V_{2,T} \times V_{2,T}$.

Now, given $f \in L^2(X \times \Omega)$, the linear form $v \mapsto (f, v)$ is a continuous functional on $V_{2,T}$, and $\overline{a}_T(u, v) := \overline{a}(Tu, Tv)$ is an inner product on $V_{2,T}$. So there exists a unique $u' \in V_{2,T}$ such that

$$(f, v) = \overline{a}_T(u', v) = \overline{a}(Tu', Tv) = a(Tu', v) \quad \forall\, v \in V_{2,T}.$$

Because T is a one-to-one map, we set $u = Tu' \in V_1$ and then

$$(f, v) = a(u, v) \quad \forall\, v \in V_{2,T}.$$

This implies that $u \in V_1$ is a weak solution of the BVP (2.2.1). If $u \in V_2$, then it is also unique. This is shown as follows. Let $\tilde{u} \in V_2$ be another weak solution. Then

$$a(u - \tilde{u}, v) = 0 \quad \forall\, v \in V_2.$$

Take $v = u - \tilde{u}$ and apply the relation (2.2.8), we conclude that $u - \tilde{u} = 0$.

We summarize the above discussion as the following theorem:

Theorem 2.2.2. *Assume* (2.2.3). *Then, for an arbitrary $f \in L^2(X \times \Omega)$, there exists $u \in V_1$ satisfying*

$$a(u, v) = (f, v) \quad \forall\, v \in V_2.$$

Moreover, if $u \in V_2$, then it is unique.

If the solution u thus obtained is of C^2 on $\overline{X} \times \Omega$, then it can be verified directly that u solves the boundary value problem

$$Au = f \quad \text{in } X \times \Omega, \qquad u = 0 \quad \text{on } \Gamma_-.$$

2.3 Maximum Principle and Positivity of Solution

In this section, we present maximum principles for the Fokker-Planck equation (2.2.1), and as a consequence, we show a positivity property of the solution. We first introduce a lemma.

Lemma 2.3.1. *Let $w \in C^2(\Omega)$. If w achieves its maximum at a point $\boldsymbol{\omega}_0 \in \Omega$, then*

$$\Delta^* w(\boldsymbol{\omega}_0) \leq 0.$$

Proof. Define $w^*(\boldsymbol{x}) = w(\boldsymbol{x}/|\boldsymbol{x}|)$ for $\boldsymbol{x} \in \mathbb{R}^3 \setminus \{0\}$. If w achieves its maximum at $\boldsymbol{\omega}_0 \in \Omega$, then w^* achieves its maximum at any $\boldsymbol{x}_0 \neq \mathbf{0}$ with $\boldsymbol{\omega}_0 = \boldsymbol{x}_0/|\boldsymbol{x}_0|$. Then by

the definition of Laplace-Beltrami operator on Ω, we have

$$\Delta^* w(\boldsymbol{\omega}_0) = \Delta w(\frac{\boldsymbol{x}_0}{|\boldsymbol{x}_0|}) = \Delta w^*(\boldsymbol{x}_0) \leq 0.$$

Therefore, the stated result holds. □

To simplify the notation, we introduce the set

$$X_\Omega = \Gamma_+ \cup \Gamma_0 \cup (X \times \Omega),$$

and define $C^{1,2}(X_\Omega)$ to be the space of all functions $v \in C(X_\Omega)$ such that $\boldsymbol{\nabla}^* v$, $\Delta^* v$ and $\boldsymbol{\nabla} v$ are all continuous on X_Ω. We have the following maximum principle.

Theorem 2.3.2. *Assume* $u \in C^{1,2}(X_\Omega) \cap C(\bar{X} \times \Omega)$ *and*

$$a_1(\boldsymbol{x}) > 0, \quad a_2(\boldsymbol{x}) = 0 \quad \textit{for } \boldsymbol{x} \in X.$$

(i) *If*

$$Au < 0 \quad \textit{in } X_\Omega, \tag{2.3.1}$$

then

$$\max_{\bar{X}\times\Omega} u = \max_{\Gamma_-} u.$$

(ii) *Likewise, if*

$$Au > 0 \quad \textit{in } X_\Omega,$$

then

$$\min_{\bar{X}\times\Omega} u = \min_{\Gamma_-} u.$$

Proof. We first prove (i). Let $(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \in X_\Omega$ be a point with $u(\boldsymbol{x}_0, \boldsymbol{\omega}_0) = \max_{\overline{X}\times\Omega} u$. By Lemma 2.3.1,

$$-a_1 \Delta^* u(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \geq 0. \tag{2.3.2}$$

We distinguish three cases according to the location of the point $(\boldsymbol{x}_0, \boldsymbol{\omega}_0)$.

First, if $(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \in X \times \Omega$, then

$$\nabla u(\boldsymbol{x}_0, \boldsymbol{\omega}_0) = 0. \tag{2.3.3}$$

Next, if $(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \in \Gamma_+$, then

$$\boldsymbol{\omega}_0 \cdot \nabla u(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \geq 0. \tag{2.3.4}$$

Indeed, the condition $(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \in \Gamma_+$ implies $\boldsymbol{\omega}_0 \cdot \boldsymbol{\nu}(\boldsymbol{x}_0) > 0$. For $t > 0$ sufficiently small,

$$u(\boldsymbol{x}_0 - t\boldsymbol{\omega}_0, \boldsymbol{\omega}_0) \leq u(\boldsymbol{x}_0, \boldsymbol{\omega}_0).$$

Rearranging and dividing by t, we get

$$\frac{u(\boldsymbol{x}_0, \boldsymbol{\omega}_0) - u(\boldsymbol{x}_0 - t\boldsymbol{\omega}_0, \boldsymbol{\omega}_0)}{t} \geq 0.$$

Letting $t \to 0^+$, we obtain (2.3.4).

Finally, suppose $(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \in \Gamma_0$. Let us show that

$$\boldsymbol{\omega}_0 \cdot \nabla u(\boldsymbol{x}_0, \boldsymbol{\omega}_0) = 0. \tag{2.3.5}$$

For this purpose, we choose a continuous boundary curve segment $\{\boldsymbol{x}(t) \mid |t| < t_0\} \subset \partial X$, $t_0 > 0$ sufficiently small, such that $\boldsymbol{x}(0) = \boldsymbol{x}_0$, and the direction $\boldsymbol{\omega}(t) := (\boldsymbol{x}(t) - \boldsymbol{x}_0)/\|\boldsymbol{x}(t) - \boldsymbol{x}_0\|$ has the properties that $\boldsymbol{\omega}(t)\cdot\boldsymbol{\omega}_0 > 0$ for $t \in (0, t_0)$ and

$\boldsymbol{\omega}(t) \to \boldsymbol{\omega}_0$ as $t \to 0+$, and $\boldsymbol{\omega}(t)\cdot\boldsymbol{\omega}_0 < 0$ for $t \in (-t_0, 0)$ and $\boldsymbol{\omega}(t) \to -\boldsymbol{\omega}_0$ as $t \to 0-$. Since u attains its maximum over $\bar{X}\times\Omega$ at $(\boldsymbol{x}_0, \boldsymbol{\omega}_0)$, we have $u(\boldsymbol{x}(t), \boldsymbol{\omega}_0) \le u(\boldsymbol{x}_0, \boldsymbol{\omega}_0)$. Thus,

$$\frac{u(\boldsymbol{x}_0 + \|\boldsymbol{x}(t) - \boldsymbol{x}_0\|\boldsymbol{\omega}(t)) - u(\boldsymbol{x}_0, \boldsymbol{\omega}_0)}{\|\boldsymbol{x}(t) - \boldsymbol{x}_0\|} \le 0.$$

Taking the limits $t \to 0+$ and $t \to 0-$, we get

$$\boldsymbol{\omega}_0 \cdot \nabla u(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \le 0, \qquad -\boldsymbol{\omega}_0 \cdot \nabla u(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \le 0,$$

respectively. Therefore, (2.3.5) holds.

Combining (2.3.2) with (2.3.3) or (2.3.4) or (2.3.5), we know that in any case, $Au(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \ge 0$, contradicting to the condition (2.3.1). Therefore, (i) holds. As $A(-u) < 0$ whenever $Au > 0$, assertion (ii) follows from (i). □

Next we include the zeroth-order term. Denote $u^+ = \max\{u, 0\}$, $u^- = \min\{u, 0\}$.

Theorem 2.3.3. *Assume* $u \in C^{1,2}(X_\Omega) \cap C(\bar{X}\times\Omega)$ *and*

$$a_1(\boldsymbol{x}) > 0, \quad a_2(\boldsymbol{x}) > 0 \quad \textit{for } \boldsymbol{x} \in X.$$

(i) *If*

$$Au \le 0 \quad \textit{in } X_\Omega, \tag{2.3.6}$$

then

$$\max_{\bar{X}\times\Omega} u \le \max_{\Gamma_-} u^+.$$

(ii) *Likewise, if*

$$Au \ge 0 \quad \textit{in } X_\Omega,$$

then

$$\min_{\bar{X}\times\Omega} u \geq \min_{\Gamma_-} u^-.$$

Proof. As in the proof of Theorem 2.3.2, we only prove (i). Thus, assume (2.3.6) and let u attain a positive maximum at a point $(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \in X_\Omega$. As in the proof of Theorem 2.3.2, we have

$$Au(\boldsymbol{x}_0, \boldsymbol{\omega}_0) \geq a_2(\boldsymbol{x}_0)\, u(\boldsymbol{x}_0, \boldsymbol{\omega}_0) > 0.$$

This contradicts to the assumption (2.3.6). □

An immediate consequence of the above theorem is the following result, which is an important property for the Fokker-Planck equation to make physical sense.

Corollary 2.3.4. *Assume* $a_1(\boldsymbol{x}) > 0$ *and* $a_2(\boldsymbol{x}) > 0$ *for* $\boldsymbol{x} \in X$. *If* $u \in C^{1,2}(X_\Omega) \cap C(\bar{X} \times \Omega)$ *satisfies* (2.2.1) *with* $f \geq 0$ *in* $X \times \Omega$, *then* $u \geq 0$ *in* X_Ω.

CHAPTER 3

DIFFERENTIAL APPROXIMATIONS AND THEIR APPLICATIONS TO INVERSE PROBLEMS

In [20], a preliminary study of a family of differential approximations of the RTE (RT/DA) was provided. In this chapter, we give a mathematical study of the RT/DA equations, as well as numerical experiments on the corresponding inverse problems.

3.1 Differential Approximation of The Integral Operator

The idea of the derivation of the RT/DA equations is based on the approximation of the integral operator S by a sequence of linear combinations of the inverse of linear elliptic differential operators on the unit sphere ([20]). The point of departure of the approach is the knowledge of eigenvalues and eigenfunctions of the operator S. Specifically, for a spherical harmonic of order n, $Y_n(\boldsymbol{\omega})$ (cf. [4] for an introduction and spherical harmonics),

$$(SY_n)(\boldsymbol{\omega}) = k_n Y_n(\boldsymbol{\omega}),\ k_n = 2\pi \int_{-1}^{1} k(s) P_n(s)\, \mathrm{d}s, \tag{3.1.1}$$

where P_n is the Legendre polynomial of deg. n. In other words, k_n is an eigenvalue of S with spherical harmonics of order n as corresponding eigenfunctions. The eigenvalues have the property that

$$\{k_n\} \text{ is bounded, and } k_n \to 0 \text{ as } n \to \infty. \tag{3.1.2}$$

Denote by Δ^* the Laplace-Beltrami operator on the unit sphere Ω. Then

$$-(\Delta^* Y_n)(\boldsymbol{\omega}) = n(n+1)Y_n(\boldsymbol{\omega}).$$

Let $\{Y_{n,m} | -n \le m \le n, n \ge 0\}$ be an orthogonal basis in $L^2\Omega$. We have the expansion

$$u(\boldsymbol{\omega}) = \sum_{n=0}^{\infty} \sum_{m=-n}^{n} u_{n,m} Y_{n,m}(\boldsymbol{\omega}) \text{ in } L^2(\Omega),\ u_{n,m} = \int_{\Omega} u(\boldsymbol{\omega}) Y_{n,m}\, \mathrm{d}\sigma(\boldsymbol{\omega}).$$

With such an expansion of $u \in L^2(\Omega)$, we have an expansion for Su:

$$Su(\boldsymbol{\omega}) = \sum_{n=0}^{\infty} k_n \sum_{m=-n}^{n} u_{n,m} Y_{n,m}(\boldsymbol{\omega}) \text{ in } L^2(\Omega).$$

Suppose there are real numbers $\{\lambda_i, \alpha_i\}_{i \ge 1}$ such that

$$\sum_{i=1}^{\infty} \frac{\lambda_i}{1 + n\,(n+1)\,\alpha_i} = k_n,\ n = 0, 1, \cdots .$$

Then formally,

$$S = \sum_{i=1}^{\infty} \lambda_i (I - \alpha_i \Delta^*)^{-1}. \tag{3.1.3}$$

The formal equality (3.1.3) motivates us to consider approximating S by

$$S_j = \sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1}, \quad j \ge 1. \tag{3.1.4}$$

The eigenvalues of S_j are $\sum_{i=1}^{j} \lambda_{j,i} (1 + n\,(n+1)\,\alpha_{j,i})^{-1}$ with associated eigenfunctions the spherical harmonics of order n:

$$(S_j Y_n)(\boldsymbol{\omega}) = \left[\sum_{i=1}^{j} \frac{\lambda_{j,i}}{1 + n\,(n+1)\,\alpha_{j,i}} \right] Y_n(\boldsymbol{\omega}). \tag{3.1.5}$$

Note that for a fixed j,

$$\sum_{i=1}^{j} \frac{\lambda_{j,i}}{1 + n\,(n+1)\,\alpha_{j,i}} \to 0 \ \text{ as } n \to \infty.$$

Thus, the eigenvalue sequence of S_j has a unique accumulation point 0, a property for the operator S, cf. (3.1.2). Hence, we choose the parameters $\{\lambda_{j,i}, \alpha_{j,i}\}_{i=1}^{j}$ so that for some integer n_j depending on j,

$$\sum_{i=1}^{j} \frac{\lambda_{j,i}}{1 + n\,(n+1)\,\alpha_{j,i}} = k_n,\ n = 0, 1, \cdots, n_j - 1. \tag{3.1.6}$$

We require $n_j \to \infty$ as $j \to \infty$.

We have (cf. [4]):

Theorem 3.1.1. *Under the assumption* (3.1.6) *and*

$$\sup_{n \geq n_j} \left| \sum_{i=1}^{j} \frac{\lambda_{j,i}}{1 + n\,(n+1)\,\alpha_{j,i}} \right| \to 0 \ \textit{as}\ j \to \infty, \tag{3.1.7}$$

we have the convergence: $\|S_j - S\|_{\mathcal{L}(L^2(\Omega), L^2(\Omega))} \to 0$ *as* $j \to \infty$.

A sufficient condition for (3.1.7) is that all $\lambda_{j,i}$ and $\alpha_{j,i}$ are positive.

Theorem 3.1.2. *Assume* (3.1.6) *and* $\lambda_{j,i} > 0$ *and* $\alpha_{j,i} > 0$ *for* $i = 1, \cdots, j$. *Then* (3.1.7) *holds.*

Notice that $\alpha_{j,i} > 0$ is needed to ensure ellipticity of the differential operator $(I - \alpha_{j,i}\Delta^*)$. When we discretize the operator S_j , the positivity of $\{\lambda_{j,i}\}_{i=1}^{j}$ is desirable for numerical stability in computing approximations of S_j.

Consider an operator S_j of the form (3.1.4) to approximate S. As noted after Theorem (3.1.2), to maintain ellipticity of the differential operator $(I - \alpha_{j,i}\Delta^*)$ and for stable numerical approximation of the operator S_j , we require

$$\alpha_{j,i} > 0,\ \lambda_{j,i} > 0,\ 1 \leq i \leq j. \tag{3.1.8}$$

Recall the property (3.1.1); for the numbers $\{k_n\}$, we assume $k_0 \geq k_1 \geq \cdots$. This assumption is quite reasonable and is valid for phase functions in practical use.

From now on, we drop the letter j in the subscripts for $\lambda_{j,i}$ and $\alpha_{j,i}$. To get some idea about the operators S_j, we consider the special cases $j = 1, 2, 3$ next. For $j = 1$, we have

$$S_1 Y_n(\boldsymbol{\omega}) = k_{1,n} Y_n(\boldsymbol{\omega}), \quad k_{1,n} = \frac{\lambda_1}{1 + \alpha_1 n\,(n+1)}.$$

Equate the first two eigenvalues of S and S_1:

$$\lambda_1 = k_0, \qquad \frac{\lambda_1}{1 + 2\,\alpha_1} = k_1.$$

Thus,

$$\lambda_1 = k_0, \quad \alpha_1 = \frac{1}{2}\left(\frac{k_0}{k_1} - 1\right). \tag{3.1.9}$$

Note that $\lambda_1 > 0$, $\alpha_1 > 0$.

For $j = 2$,

$$S_2 = \lambda_1 (I - \alpha_1 \Delta^*)^{-1} + \lambda_2 (I - \alpha_2 \Delta^*)^{-1},$$

$$\lambda_1 > 0, \quad \lambda_2 > 0, \quad \alpha_1 > 0, \quad \alpha_2 > 0, \quad \alpha_1 \neq \alpha_2.$$

We have

$$S_2 Y_n(\boldsymbol{\omega}) = k_{2,n} Y_n(\boldsymbol{\omega}),\ k_{2,n} = \frac{\lambda_1}{1 + \alpha_1 n\,(n+1)} + \frac{\lambda_2}{1 + \alpha_2 n\,(n+1)}.$$

Require the parameters to match the first three eigenvalues $k_{2,0} = k_0,\ k_{2,1} = k_1,\ k_{2,2} =$

k_2, i.e.,

$$\lambda_1 + \lambda_2 = k_0, \tag{3.1.10}$$

$$\frac{\lambda_1}{1+2\alpha_1} + \frac{\lambda_2}{1+2\alpha_2} = k_1, \tag{3.1.11}$$

$$\frac{\lambda_1}{1+6\alpha_1} + \frac{\lambda_2}{1+6\alpha_2} = k_2. \tag{3.1.12}$$

Use α_1 as the parameter for the solution of the resulting system of three equations. It is shown in [20] that

$$\begin{aligned}
\alpha_2 &= \frac{1}{6} \cdot \frac{(3\,k_1 - 2\,k_0 - k_2) + 6\,(k_1 - k_2)\,\alpha_1}{(k_2 - k_1) + 2\,(3\,k_2 - k_1)\,\alpha_1}, \\
\lambda_2 &= \frac{2\,[(k_1 - k_0) + 2\,k_1\alpha_1]\,[(k_2 - k_0) + 6\,k_2\alpha_1]}{(2\,k_0 + k_2 - 3\,k_1) + 12\,(k_2 - k_1)\,\alpha_1 + 12\,(3\,k_2 - k_1)\,\alpha_1^2}, \\
\lambda_1 &= 1 - \lambda_2.
\end{aligned}$$

The issue of positivity of the solution $(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$ is also discussed in [20].

For $j = 3$,

$$S_3 = \lambda_1(I - \alpha_1\Delta^*)^{-1} + \lambda_2(I - \alpha_2\Delta^*)^{-1} + \lambda_3(I - \alpha_3\Delta^*)^{-1} \tag{3.1.13}$$

with the parameters satisfying

$$\lambda_1 > 0, \quad \lambda_2 > 0, \quad \lambda_3 > 0, \quad \alpha_1 > 0, \quad \alpha_2 > 0, \quad \alpha_3 > 0, \qquad \alpha_1 \neq \alpha_2 \neq \alpha_3. \tag{3.1.14}$$

We have

$$S_3 Y_n(\boldsymbol{\omega}) = k_{3,n} Y_n(\boldsymbol{\omega}), \tag{3.1.15}$$

and

$$k_{3,n} = \frac{\lambda_1}{1 + \alpha_1 n(n+1)} + \frac{\lambda_2}{1 + \alpha_2 n(n+1)} + \frac{\lambda_3}{1 + \alpha_3 n(n+1)}. \tag{3.1.16}$$

Require the parameters to match the first four eigenvalues

$$k_{3,0} = k_0, \quad k_{3,1} = k_1, \quad k_{3,2} = k_2, \quad k_{3,3} = k_3, \tag{3.1.17}$$

i. e. ,

$$\lambda_1 + \lambda_2 + \lambda_3 = k_0, \tag{3.1.18}$$

$$\frac{\lambda_1}{1+2\alpha_1} + \frac{\lambda_2}{1+2\alpha_2} + \frac{\lambda_3}{1+2\alpha_3} = k_1, \tag{3.1.19}$$

$$\frac{\lambda_1}{1+6\alpha_1} + \frac{\lambda_2}{1+6\alpha_2} + \frac{\lambda_3}{1+6\alpha_3} = k_2, \tag{3.1.20}$$

$$\frac{\lambda_1}{1+12\alpha_1} + \frac{\lambda_2}{1+12\alpha_2} + \frac{\lambda_3}{1+12\alpha_3} = k_3. \tag{3.1.21}$$

Consider the system (3.1.18)-(3.1.21) for a general form solution. Using α_2 and α_3 as the parameter for the solution, we get

$$\begin{aligned}\lambda_1 = (&5(2k_0 - 3k_1(1+2\alpha_2)(1+2\alpha_3) + k_2(1+6\alpha_2)(1+6\alpha_3))\\ &(5k_0 - 6k_1(1+2\alpha_2)(1+2\alpha_3) +\\ &k_3(1+12\alpha_2)(1+12\alpha_3))(k_0 - 2k_2(1+6\alpha_2)(1+6\alpha_3) + k_3(1+12\alpha_2)(1+12\alpha_3)))/\\ &(2(5k_0 - 9k_1(1+2\alpha_2)^2(1+2\alpha_3) + 5k_2(1+6\alpha_2)^2(1+6\alpha_3) - k_3(1+12\alpha_2)^2(1+12\alpha_3))\\ &(5k_0 - 9k_1(1+2\alpha_2)(1+2\alpha_3)^2 + 5k_2(1+6\alpha_2)(1+6\alpha_3)^2 - k_3(1+12\alpha_2)(1+12\alpha_3)^2)),\end{aligned} \tag{3.1.22}$$

$$\begin{aligned}\lambda_2 = (&(1+2\alpha_2)(1+6\alpha_2)(1+12\alpha_2)(k_2k_3(1+6\alpha_3)(1+12\alpha_3) +\\ &k_0(k_1 - 5k_2 + 4k_3 + 2(k_1 - 15k_2 + 24k_3)\alpha_3)\\ &+k_1(1+2\alpha_3)(4k_2(1+6\alpha_3) - 5k_3(1+12\alpha_3)))) /\\ &\big(4(\alpha_2 - \alpha_3)\big(-5k_0 + 9k_1(1+2\alpha_2)^2(1+2\alpha_3) - 5k_2(1+6\alpha_2)^2(1+6\alpha_3)\\ &+ k_3(1+12\alpha_2)^2(1+12\alpha_3)\big)\big),\end{aligned} \tag{3.1.23}$$

$$\lambda_3 = ((k_2k_3\,(1+6\alpha_2)\,(1+12\alpha_2) + k_0\,(k_1 - 5k_2 + 4k_3 + 2\,(k_1 - 15k_2 + 24k_3)\,\alpha_2) +$$

$$k_1\,(1+2\alpha_2)\,(4k_2\,(1+6\alpha_2) - 5k_3\,(1+12\alpha_2)))\,(1+2\alpha_3)\,(1+6\alpha_3)\,(1+12\alpha_3)) \,/$$

$$\big(4\,(\alpha_2-\alpha_3)\,\big(-5k_0 + 9k_1\,(1+2\alpha_2)\,(1+2\alpha_3)^{\,2} - 5k_2\,(1+6\alpha_2)\,(1+6\alpha_3)^{\,2}$$

$$+\ k_3\,(1+12\alpha_2)\,(1+12\alpha_3)^{\,2}\big)\big)\,, \quad (3.1.24)$$

$$\alpha_1 =$$

$$\frac{5k_0 - 9k_1\,(1+2\alpha_2)\,(1+2\alpha_3) + 5k_2\,(1+6\alpha_2)\,(1+6\alpha_3) - k_3\,(1+12\alpha_2)\,(1+12\alpha_3)}{6\,(3k_1\,(1+2\alpha_2)\,(1+2\alpha_3) - 5k_2\,(1+6\alpha_2)\,(1+6\alpha_3) + 2k_3\,(1+12\alpha_2)\,(1+12\alpha_3))}. \quad (3.1.25)$$

Next, we use the Henyey-Greenstein phase function as an example; in this case,

$$k_n = g^n,\ n = 0, 1, \cdots .$$

For the one term approximation $S_1 = \lambda_1(I - \alpha_1\Delta^*)^{-1}$, from (3.1.9) we have

$$\lambda_1 = k_0, \quad \alpha_1 = \frac{1-g}{2g}. \quad (3.1.26)$$

For the two term approximation $S_2 = \lambda_1(I - \alpha_1\Delta^*)^{-1} + \lambda_2(I - \alpha_2\Delta^*)^{-1}$, we have

$$\lambda_1 = \frac{g\,(1-g)\,(2g-1)\,(1+8\alpha_1+12\alpha_1^2)}{(1-g)\,(2-g) + 12g\,(g-1)\,\alpha_1 + 12g\,(3g-1)\,\alpha_1^2},$$
$$\lambda_2 = \frac{2\,(g-1+2g\alpha_1)\,(g^2-1+6g^2\alpha_1)}{(1-g)\,(2-g) + 12g\,(g-1)\,\alpha_1 + 12g\,(3g-1)\,\alpha_1^2},$$
$$\alpha_2 = \frac{1-g}{6g}\cdot\frac{g-2+6\,g\,\alpha_1}{g-1+2\,(3g-1)\,\alpha_1}.$$

On the issue of positivity of the one parameter solution $(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$, with $\alpha_1 > 0$, it is shown in [20] that under the assumption $g > 1 = 2$, valid in applications with

highly forward peaked scattering, the condition for a positive solution $(\alpha_1, \alpha_2, \lambda_1, \lambda_2)$ is

$$\alpha_1 > \frac{2-g}{6g}. \tag{3.1.27}$$

Since $\alpha_1 = 1/2$ satisfies (3.1.27), one solution is

$$\alpha_1 = \frac{1}{2},\ \alpha_2 = \frac{1-g}{6g},\ \lambda_1 = \frac{4g(1-g)}{4g-1},\ \lambda_2 = \frac{4g^2-1}{4g-1}. \tag{3.1.28}$$

Now consider $j = 3$,

$$\lambda_1 + \lambda_2 + \lambda_3 = 1, \tag{3.1.29}$$

$$\frac{\lambda_1}{1+2\alpha_1} + \frac{\lambda_2}{1+2\alpha_2} + \frac{\lambda_3}{1+2\alpha_3} = g, \tag{3.1.30}$$

$$\frac{\lambda_1}{1+6\alpha_1} + \frac{\lambda_2}{1+6\alpha_2} + \frac{\lambda_3}{1+6\alpha_3} = g^2, \tag{3.1.31}$$

$$\frac{\lambda_1}{1+12\alpha_1} + \frac{\lambda_2}{1+12\alpha_2} + \frac{\lambda_3}{1+12\alpha_3} = g^3. \tag{3.1.32}$$

We choose α_1 and α_2, positive and distinct, as the parameters and express the other quantities in terms of them. There are many positive solution sets to the system (3.1.29)-(3.1.32) with positive parameters α_1 and α_2. For the numerical examples, we use parameter sets so that overall the eigenvalues of S_3 are close to those of S. In particular, for $g = 0.9$, we choose

$$\alpha_1 = 0.00957621, \quad \alpha_2 = 0.08, \quad \alpha_3 = 0.712,$$

$$\lambda_1 = 0.660947, \quad \lambda_2 = 0.248262, \quad \lambda_3 = 0.0907913;$$

For $g = 0.95$, we choose

$$\alpha_1 = 0.00325598, \quad \alpha_2 = 0.06, \quad \alpha_3 = 0.701,$$

$$\lambda_1 = 0.78042, \quad \lambda_2 = 0.174622, \quad \lambda_3 = 0.0449584;$$

and for $g = 0.99$, we choose

$$\alpha_1 = 0.000306188, \quad \alpha_2 = 0.05, \quad \alpha_3 = 0.95,$$

$$\lambda_1 = 0.940247, \quad \lambda_2 = 0.0526772, \quad \lambda_3 = 0.00707558.$$

For $g = 0.9$, we compare the eigenvalues of S_j for $j = 1, 2, 3$ with those of S in Figure 3.1, 3.2, and 3.3, respectively. From these figures, we can tell that the approximation of S_3 should be more accurate than that of S_2, which should be in turn more accurate than S_1. This observation is valid for other values of g below.

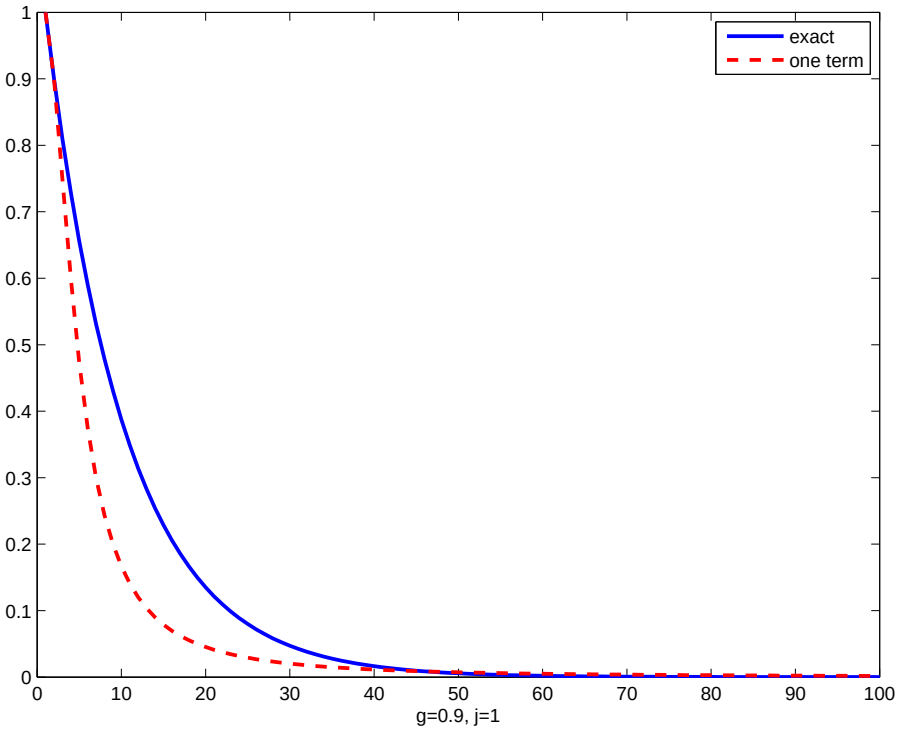

Figure 3.1. One-term approximation for $g = 0.9$

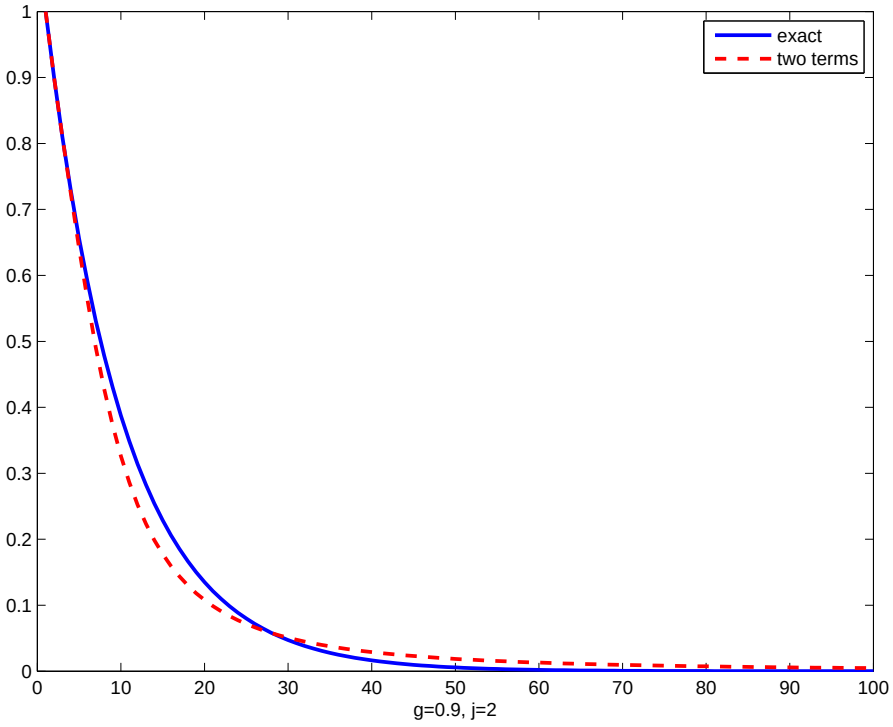

Figure 3.2. Two-term approximation for $g = 0.9$

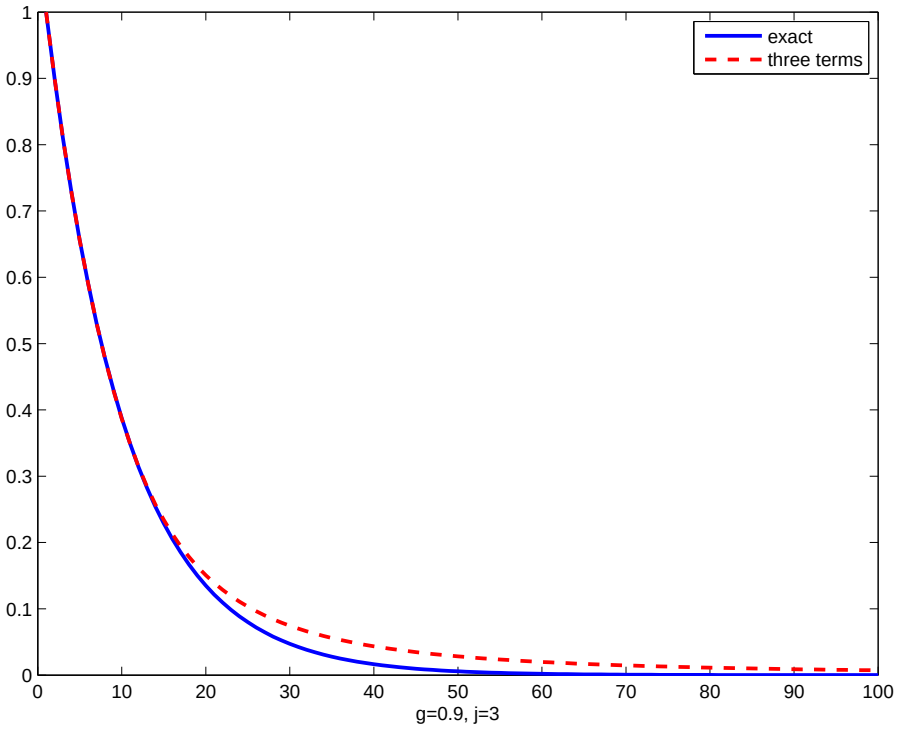

Figure 3.3. Three-term approximation for $g = 0.9$

3.1.1 An Optimization Problem for Identifying the Parameters of Three Terms Approximation

From the above discussion we know that when looking for the parameters of $\mathrm{RT}/\mathrm{DA}_j$, there are usually more unknowns than the number of equations. To solve the equations uniquely, we treat several unknown parameter as fixed variables, which creates infinitely many sets of solutions. This leaves us a huge freedom to identify the best set of parameters among all possible choices. To address this problem, we apply optimization techniques.

Define the object function

$$f(\alpha_2, \alpha_3) = \sum_{n=4}^{50} \left| \frac{\lambda_1}{1 + n(n+1)\alpha_1} + \frac{\lambda_2}{1 + n(n+1)\alpha_2} + \frac{\lambda_3}{1 + n(n+1)\alpha_3} - g^n \right|^2,$$

where λ_1, λ_2, λ_3 and α_1 are determined by the equations:

$$\lambda_1 + \lambda_2 + \lambda_3 = 1, \tag{3.1.33}$$

$$\frac{\lambda_1}{1 + 2\alpha_1} + \frac{\lambda_2}{1 + 2\alpha_2} + \frac{\lambda_3}{1 + 2\alpha_3} = g, \tag{3.1.34}$$

$$\frac{\lambda_1}{1 + 6\alpha_1} + \frac{\lambda_2}{1 + 6\alpha_2} + \frac{\lambda_3}{1 + 6\alpha_3} = g^2, \tag{3.1.35}$$

$$\frac{\lambda_1}{1 + 12\alpha_1} + \frac{\lambda_2}{1 + 12\alpha_2} + \frac{\lambda_3}{1 + 12\alpha_3} = g^3. \tag{3.1.36}$$

Consider the following optimization problem:

$$\underset{\alpha_2, \alpha_3 > 0}{\operatorname{argmin}} f(\alpha_2, \alpha_3). \tag{3.1.37}$$

Notice that in the definition of the object function f, we choose $n = 50$ for the upper bound of the summation. For the values of g considered in the our applications (e.g. $g = 0.9$ or 0.95), g^{50} is already quite close to 0. Since we have

$\left|\sum_{j=1}^{3} \frac{\lambda_j}{1+n(n+1)\alpha_j} - g^n\right| < \max(\sum_{j=1}^{3} \frac{\lambda_j}{1+n(n+1)\alpha_j}, g^n)$ for all positive α_j and λ_j, including terms with large value of n only slightly changes the value of the object function f while it greatly increases the computation complexity. However, if g is even further close to 1, say $g = 0.99$, greater value of n should be used so that higher number of terms would be included in the summation, which results in a better set of parameters.

Since the function f is non-convex, a global minimizer is difficult to find. To assist solving the minimization problem, we graph the function f as in Fig. 3.4 to visually identify good initial values for the algorithm to reach a global minimizer.

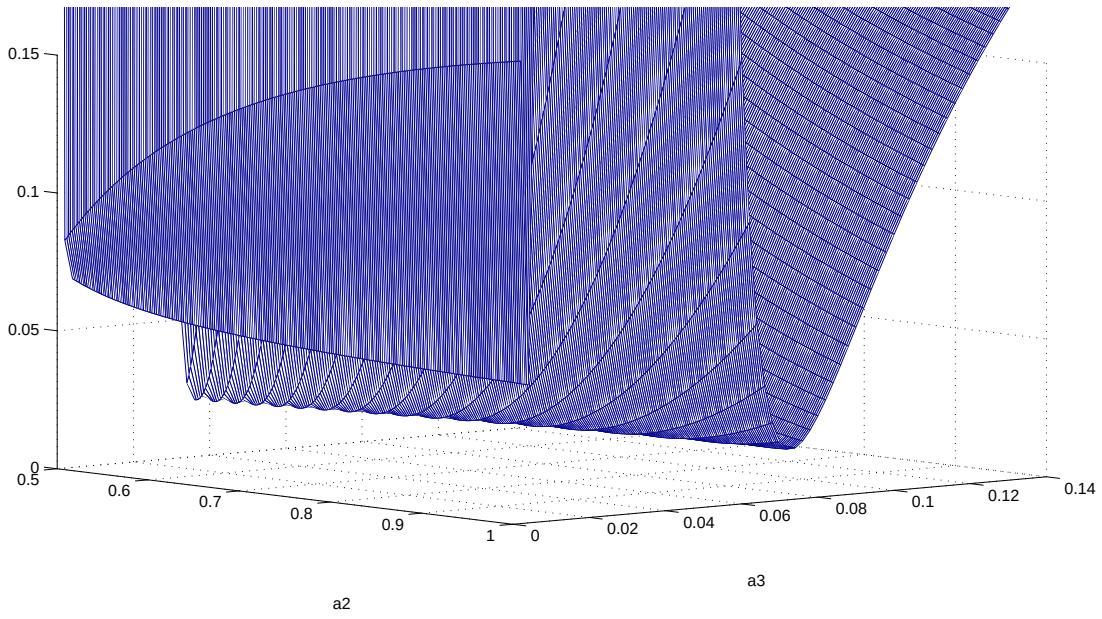

Figure 3.4. The values of object function in terms of α_2 and α_3 (denoted by $a2$ and $a3$ in the Figure) for $g = 0.95$

For $g = 0.9$, the numerical solution to this problem is

$$\lambda_1 = 0.766383864642871, \lambda_2 = 0.068551521590490, \lambda_3 = 0.165064613766639,$$

$$\alpha_1 = 0.012970044697335, \alpha_2 = 0.846476840302828, \alpha_3 = 0.147121050904354.$$

For $g = 0.95$, the numerical solution to this problem is

$$\lambda_1 = 0.789349144823840, \lambda_2 = 0.176816745780522, \lambda_3 = 0.033834109395638,$$

$$\alpha_1 = 0.003171611003095, \alpha_2 = 0.072819590072973, \alpha_3 = 0.998760488158562.$$

3.2 Well Posedness of RT/DA Problems

We use S_j of (3.1.4) for the approximation of the integral operator S. In the following, we drop the subscript j in the parameters $\lambda_{j,i}$ and $\alpha_{j,i}$ for S_j and write

$$S_j u(\boldsymbol{x}, \boldsymbol{\omega}) = \sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1} u(\boldsymbol{x}, \boldsymbol{\omega}), \quad j \geq 1.$$

Then the RT/DA$_j$ problem is

$$\boldsymbol{\omega} \cdot \nabla u(\boldsymbol{x}, \boldsymbol{\omega}) + \mu_t u(\boldsymbol{x}, \boldsymbol{\omega}) = \mu_s(\boldsymbol{x}) S_j u(\boldsymbol{x}, \boldsymbol{\omega}) + f(\boldsymbol{x}, \boldsymbol{\omega}), \ (\boldsymbol{x}, \boldsymbol{\omega}) \in X \times \Omega, \tag{3.2.1}$$

$$u(\boldsymbol{x}, \boldsymbol{\omega}) = u_{in}(\boldsymbol{x}, \boldsymbol{\omega}), \ (\boldsymbol{x}, \boldsymbol{\omega}) \in \Gamma_-. \tag{3.2.2}$$

Let us consider the well-posedness of (3.2.1)–(3.2.2). Introduce

$$w_i(\boldsymbol{x}, \boldsymbol{\omega}) = (I - \alpha_i \Delta^*)^{-1} u(\boldsymbol{x}, \boldsymbol{\omega}), \tag{3.2.3}$$

$$w(\boldsymbol{x}, \boldsymbol{\omega}) = \sum_{i=1}^{j} \lambda_i w_i(\boldsymbol{x}, \boldsymbol{\omega}). \tag{3.2.4}$$

Then the equation (3.2.1) can be rewritten as

$$\boldsymbol{\omega} \cdot \nabla u(\boldsymbol{x}, \boldsymbol{\omega}) + \mu_t u(\boldsymbol{x}, \boldsymbol{\omega}) = \mu_s(\boldsymbol{x}) w(\boldsymbol{x}, \boldsymbol{\omega}) + f(\boldsymbol{x}, \boldsymbol{\omega}), \ (\boldsymbol{x}, \boldsymbol{\omega}) \in X \times \Omega. \tag{3.2.5}$$

For simplicity we limit the analysis to the case where X is a convex domain in $\mathbb{R}^3$. The argument can be extended to a domain X satisfying the generalized convexity condition without problem ([1]). Then for each $\boldsymbol{\omega} \in \Omega$ and each $\boldsymbol{z} \in X_{\boldsymbol{\omega}}$, $X_{\boldsymbol{\omega},\boldsymbol{x}}$ is the line segment

$$X_{\boldsymbol{\omega},\boldsymbol{z}} = \{\boldsymbol{z} + s\,\boldsymbol{\omega} | s \in (s_-, s_+)\},$$

where $s_\pm = s_\pm(\boldsymbol{\omega}, \boldsymbol{z})$ depend on $\boldsymbol{\omega}$ and $\boldsymbol{z}$, and $\boldsymbol{x}_\pm := \boldsymbol{z} + s_\pm\boldsymbol{\omega}$ are the intersection points of the line $\{\boldsymbol{z} + s\,\boldsymbol{\omega} | s \in \mathbb{R}\}$ with ∂X.

In the following, we write $s_\pm$ instead of $s_\pm(\boldsymbol{\omega}, \boldsymbol{z})$ whenever there is no danger for confusion. We write the equation (3.2.5) as

$$\frac{\partial}{\partial s} u(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega}) + \mu_t(\boldsymbol{z} + s\boldsymbol{\omega}) u(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega}) = \mu_s(\boldsymbol{z} + s\boldsymbol{\omega}) w(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega}) + f(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega})$$

and multiply it by $\exp(\int_{s_-}^{s} \mu_t(\boldsymbol{z} + s\boldsymbol{\omega})\,\mathrm{d}s)$ to obtain

$$\begin{aligned} &\frac{\partial}{\partial s}\left(e^{\int_{s_-}^{s} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s} u(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega})\right) \\ &\qquad = e^{\int_{s_-}^{s} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s} (\mu_s(\boldsymbol{z} + s\boldsymbol{\omega}) w(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega}) + f(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega})). \end{aligned}$$

Integrate this equation from s_- to s:

$$\begin{aligned} &e^{\int_{s_-}^{s} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s} u(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega}) - u_{in} \\ &\qquad = \int_{s_-}^{s} e^{\int_{s_-}^{t} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s} (\mu_s(\boldsymbol{z} + s\boldsymbol{\omega}) w(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega}) + f(\boldsymbol{z} + s\boldsymbol{\omega}, \boldsymbol{\omega}))\,\mathrm{d}t. \end{aligned}$$

Thus, (3.2.1)–(3.2.2) is converted to a fixed-point problem

$$u = Au + F, \tag{3.2.6}$$

where

$$\begin{aligned} Au(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega}) =& \int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\mu_s(\boldsymbol{z}+s\boldsymbol{\omega})w(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})\,\mathrm{d}t \\ F(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega}) =& e^{-\int_{s_-}^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}u_{in}(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega}) \\ &+ \int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}f(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega}))\,\mathrm{d}t. \end{aligned}$$

We will show that A is a contractive mapping in a weighted $L^2(X\times\Omega)$ space. Denote $\kappa = \sup\{\mu_s(\boldsymbol{x})/\mu_t(\boldsymbol{x})|\boldsymbol{x}\in X\}$. By (1.3.5), we know that $\kappa<1$. Consider

$$\begin{aligned} &\int_{s_-}^{s_+} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})|Au(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})|^2\,\mathrm{d}s \\ &\quad= \int_{s_-}^{s_+} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\left|\int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\mu_s(\boldsymbol{z}+s\boldsymbol{\omega})w(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})\,\mathrm{d}t\right|^2\,\mathrm{d}s \\ &\qquad\le \int_{s_-}^{s_+} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\left(\int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\mu_s(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}t\right) \\ &\qquad\qquad\cdot\left(\int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\mu_s(\boldsymbol{z}+s\boldsymbol{\omega})|w(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})|^2\,\mathrm{d}t\right)^2\,\mathrm{d}s. \end{aligned}$$

Since

$$\begin{aligned} \int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\mu_s(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}t \le& \kappa\int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}t \\ =& \kappa\left(1-e^{-\int_{s_-}^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\right)\le\kappa, \end{aligned}$$

we have

$$\begin{aligned} &\int_{s_-}^{s_+} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})|Au(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})|^2\,\mathrm{d}s \\ &\quad\le \kappa\int_{s_-}^{s_+} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\mu_s(\boldsymbol{z}+s\boldsymbol{\omega})|w(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})|^2\,\mathrm{d}t\,\mathrm{d}s \\ &\quad= \kappa\int_{s_-}^{s_+} \mu_s(\boldsymbol{z}+s\boldsymbol{\omega})|w(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})|^2\left(\int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}t\right)\,\mathrm{d}s. \end{aligned}$$

Now

$$\int_{s_-}^{s} e^{-\int_t^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}t = \kappa\left(1-e^{-\int_{s_-}^s \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})\,\mathrm{d}s}\right)\le\kappa,$$

we obtain

$$\begin{aligned}\int_{s_-}^{s_+} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})|Au(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})|^2\,\mathrm{d}s \leq& \kappa\int_{s_-}^{s_+} \mu_s(\boldsymbol{z}+s\boldsymbol{\omega})|w(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})|^2\,\mathrm{d}s\\ \leq& \kappa^2\int_{s_-}^{s_+} \mu_t(\boldsymbol{z}+s\boldsymbol{\omega})|w(\boldsymbol{z}+s\boldsymbol{\omega},\boldsymbol{\omega})|^2\,\mathrm{d}s.\end{aligned}$$

Thus, we have proved the inequality

$$\|\mu_t^{1/2}Au\|_{L^2(X\times\Omega)} \leq \kappa\|\mu_t^{1/2}w\|_{L^2(X\times\Omega)}. \tag{3.2.7}$$

Returning to the definition (3.2.3), we have equivalently,

$$(I-\alpha_{j,i}\Delta^*)w_i = u \quad \text{ in } X\times\Omega.$$

For a.e. $\boldsymbol{x}\in X$, $w_i(\boldsymbol{x},\cdot)\in H^1(\Omega)$ and

$$\int_\Omega (w_i v + \alpha_i\nabla^* w_i\cdot\nabla^* v)\,\mathrm{d}\sigma(\boldsymbol{\omega}) = \int_\Omega uv\,\mathrm{d}\sigma(\boldsymbol{\omega}) \quad \forall v\in H^1(\Omega). \tag{3.2.8}$$

Since $\alpha_i>0$, this problem has a unique solution by the Lax-Milgram Lemma. Take $v(\boldsymbol{\omega}) = w_i(\boldsymbol{x},\boldsymbol{\omega})$ in (3.2.8),

$$\int_\Omega (|w_i|^2+\alpha_i|\nabla^* w_i|^2)\,\mathrm{d}\sigma(\boldsymbol{\omega}) = \int_\Omega uw_i\,\mathrm{d}\sigma(\boldsymbol{\omega}).$$

Thus,

$$\int_\Omega (|w_i|^2+2\alpha_i|\nabla^* w_i|^2)\,\mathrm{d}\sigma(\boldsymbol{\omega}) \leq \int_\Omega |u|^2\,\mathrm{d}\sigma(\boldsymbol{\omega}). \tag{3.2.9}$$

In particular,

$$\int_\Omega |w_i|^2\,\mathrm{d}\sigma(\boldsymbol{\omega}) \leq \int_\Omega |u|^2\,\mathrm{d}\sigma(\boldsymbol{\omega}).$$

Therefore

$$\|\mu_t^{1/2}w_i\|_{L^2(X\times\Omega)} \leq \kappa\|\mu_t^{1/2}u\|_{L^2(X\times\Omega)}. \tag{3.2.10}$$

Since $\lambda_i > 0$ and $\sum_{i=1}^{j} \lambda_i = 1$, from the definition (3.2.4) and (3.2.10), we get

$$\|\mu_t^{1/2} w\|_{L^2(X\times\Omega)} \leq \kappa \|\mu_t^{1/2} u\|_{L^2(X\times\Omega)}. \tag{3.2.11}$$

Combining (3.2.7) and (3.2.11), we see that the operator A: $L^2(X \times \Omega) \to L^2(X \times \Omega)$ is contractive with respect to the weighted norm $\|\mu_t^{1/2} v\|_{L^2(X\times\Omega)}$:

$$\|\mu_t^{1/2} Au\|_{L^2(X\times\Omega)} \leq \kappa \|\mu_t^{1/2} u\|_{L^2(X\times\Omega)}. \tag{3.2.12}$$

Notice that from the assumptions on μ_t, the weighted norm $\|\mu_t^{1/2} \cdot\|_{L^2(X\times\Omega)}$ is equivalent to the regular norm $\|\cdot\|_{L^2(X\times\Omega)}$. By an application of the Banach fixed-point theorem, we conclude that the equation (3.2.6) has a unique solution $u \in L^2(X \times \Omega)$. By the equation (3.2.5), we also have $\boldsymbol{\omega} \cdot \nabla u(\boldsymbol{x}, \boldsymbol{\omega}) \in L^2(X \times \Omega)$. Therefore, the solution $u \in H_2^1(X \times \Omega)$. Using (3.2.12)

$$\begin{aligned}\|\mu_t^{1/2} u\|_{L^2(X\times\Omega)} &\leq \|\mu_t^{1/2} Au\|_{L^2(X\times\Omega)} + \|\mu_t^{1/2} F\|_{L^2(X\times\Omega)} \\ &\leq \kappa \|\mu_t^{1/2} u\|_{L^2(X\times\Omega)} + c[\|u_{in}\|_{L^2(\Gamma_-)} + \|f\|_{L^2(X\times\Omega)}].\end{aligned}$$

Hence we have the bound

$$\|u\|_{L^2(X\times\Omega)} \leq c[\|u_{in}\|_{L^2(\Gamma_-)} + \|f\|_{L^2(X\times\Omega)}]. \tag{3.2.13}$$

Using (3.2.5), we obtain

$$\boldsymbol{\omega} \cdot \nabla u(\boldsymbol{x}, \boldsymbol{\omega}) = \mu_s w - \mu_t u + f.$$

Then we can improve the bound (3.2.13) to the following

$$\|u\|_{H_1^2(X\times\Omega)} \leq c[\|u_{in}\|_{L^2(\Gamma_-)} + \|f\|_{L^2(X\times\Omega)}]. \tag{3.2.14}$$

In summary, we have shown the following existence and uniqueness result.

Theorem 3.2.1. *Under the assumptions* (1.3.4), (1.3.5), (1.3.6) *and* (3.1.8), *the problem* (3.2.1)–(3.2.2) *has a unique solution* $u \in H_2^1(X \times \Omega)$, *and this solution is Lipschitz continuous with respect to the source function* f *and the boundary condition* u_{in}*: Given* $f_1, f_2 \in L^2(X \times \Omega)$ *and* $u_{in,1}, u_{in,2} \in L^2(\Gamma_-)$, *for* $k = 1,, 2$, *if* $u_k = u(f_k, u_{in,k}) \in H_1^2(X \times \Omega)$ *is the solution of the problem* (3.2.1)–(3.2.2) *with* $f = f_k$ *and* $u_{in} = u_{in,k}$, *then for some constant* c *depending only on* μ_t, μ_s, *and* X *we have the bound*

$$\|u_1 - u_2\|_{H_1^2(X\times\Omega)} \leq c[\|u_{in,1} - u_{in,2}\|_{L^2(\Gamma_-)} + \|f_1 - f_2\|_{L^2(X\times\Omega)}]. \tag{3.2.15}$$

3.3 An Iteration Method

We introduce an iteration method for solving the problem (3.2.1)–(3.2.2). Let $w^{(0)}$ be an initial guess. Then define $u^{(n)}$ and $w^{(n)}$ as follows:

$$\boldsymbol{\omega}\cdot\boldsymbol{\nabla} u^{(n)} + \mu_t\, u^{(n)} = \mu_s\, w^{(n-1)} + f \quad \text{in } X \times \Omega, \tag{3.3.1}$$

$$u^{(n)} = u_{in} \quad \text{on } \Gamma_-, \tag{3.3.2}$$

$$w_i^{(n)} = (I - \alpha_i \Delta^*)^{-1} u^{(n)}, \tag{3.3.3}$$

$$w^{(n)} = \sum_{i=1}^{j} \lambda_i w_i^{(n)}. \tag{3.3.4}$$

Denote the iteration errors $e_u^{(n)} := u - u^{(n)}$, $e_w^{(n)} := w - w^{(n)}$. Then we have the error relations

$$\begin{aligned}
\boldsymbol{\omega}\cdot\boldsymbol{\nabla} e_u^{(n)} + \mu_t\, e_u^{(n)} &= \mu_s\, e_w^{(n-1)} + f \quad \text{in } X\times\Omega,\\
e_u^{(n)} &= 0 \quad \text{on } \Gamma_-,\\
e_{w_i}^{(n)} &= (I - \alpha_i\Delta^*)^{-1} e_u^{(n)},\\
e_w^{(n)} &= \sum_{i=1}^{j} \lambda_i e_{w_i}^{(n)}.
\end{aligned}$$

We have

$$\|\sigma_t^{1/2} e_u^{(n)}\|_{L^2(X\times\Omega)} \le \kappa \|\sigma_t^{1/2} e_w^{(n)}\|_{L^2(X\times\Omega)},$$

$$\|\sigma_t^{1/2} e_w^{(n-1)}\|_{L^2(X\times\Omega)} \le \kappa \|\sigma_t^{1/2} e_u^{(n-1)}\|_{L^2(X\times\Omega)}.$$

Thus,

$$\|\sigma_t^{1/2} e_u^{(n)}\|_{L^2(X\times\Omega)} \le \kappa \|\sigma_t^{1/2} e_u^{(n-1)}\|_{L^2(X\times\Omega)},$$

and so we have

$$\|\sigma_t^{1/2} e_u^{(n)}\|_{L^2(X\times\Omega)} \le \kappa^n \|\sigma_t^{1/2} e_u^{(0)}\|_{L^2(X\times\Omega)} \to 0 \text{ as } n\to\infty.$$

Moreover, we also have the convergence of the sequence $\{w^{(n)}\}$:

$$\|\sigma_t^{1/2} e_w^{(n)}\|_{L^2(X\times\Omega)} \le \kappa \|\sigma_t^{1/2} e_u^{(n)}\|_{L^2(X\times\Omega)} \to 0 \text{ as } n\to\infty.$$

3.4 Discretizations

In this section we describe a numerical method for solving the RD/DA (3.2.1). Consider the discretization of the RT/DA equations of

$$\boldsymbol{\omega}\cdot\boldsymbol{\nabla}u + \mu_t\, u = \mu_s\, w + f, \tag{3.4.1}$$

$$u = u_{in}, \tag{3.4.2}$$

$$(I - \alpha_j\Delta^*)w_j = u, \tag{3.4.3}$$

$$w = \sum_{j=1}^{J} \lambda_j w_j. \tag{3.4.4}$$

Since (3.4.1) is essentially a hyperbolic type equation, it is natural to solve the equation by the discontinuous Galerkin (DG) discretization in space. For the elliptic equation (3.4.3), we use the regular continuous finite element method for the angular variable.

Next we give a brief description of these two numerical methods for the corresponding equations. Choose a set of nodes $\{\boldsymbol{\omega}_l\}_{l=1}^{L}$ on the unit sphere Ω. Let W^h be a finite element space with a nodal basis $\{\phi_l^{\omega}(\boldsymbol{\omega})\}_{l=1}^{L}$ corresponding to the nodes $\{\boldsymbol{\omega}_l\}_{l=1}^{L}$. By a nodal basis we refer to the property $\phi_l^{\omega}(\boldsymbol{\omega}) = \delta_{lm}$ for $1 \le l, m \le L$. Let $\overline{X} = \cup_{K\in\mathcal{T}^h} K$ be a finite element partition of $\overline{X}$, and let U^h be a corresponding finite element space. Denote by $\{\phi_i^x(\boldsymbol{x})\}_{i=1}^{I}$ a basis of U^h. The numerical solution of the equations (3.4.1)–(3.4.4) is expressed as follows:

$$u^h(\boldsymbol{x},\boldsymbol{\omega}) = \sum_{i,l} u_{il}^h \phi_i^x(\boldsymbol{x})\phi_l^{\omega}(\boldsymbol{\omega}), \qquad w^h(\boldsymbol{x},\boldsymbol{\omega}) = \sum_{i,l} w_{il}^h \phi_i^x(\boldsymbol{x})\phi_l^{\omega}(\boldsymbol{\omega}) \tag{3.4.5}$$

where $\sum_{i,l}$ stands for $\sum_{i=1}^{I}\sum_{l=1}^{L}$. For convenience, we will use the following notation:

$$u_l^h(\boldsymbol{x}) = \sum_i u_{il}^h \phi_i^x(\boldsymbol{x}), \quad u_l^h(\boldsymbol{\omega}) = \sum_l u_{il}^h \phi_i^{\omega}(\boldsymbol{\omega})$$

$$w_l^h(\boldsymbol{x}) = \sum_i w_{il}^h \phi_i^x(\boldsymbol{x}), \quad w_l^h(\boldsymbol{\omega}) = \sum_l w_{il}^h \phi_i^\omega(\boldsymbol{\omega})$$

We adopt the discrete-ordinate discontinuous Galerkin method developed in [23] for (3.4.1) and (3.4.2), and the resulting numerical formulation is given by: For $1 \le l \le L$,

$$\int_{\partial K} \widehat{u_l^h} \boldsymbol{\omega}_l \cdot \nu_K \phi_j^x \, \mathrm{d}\sigma(\boldsymbol{x}) - \int_K u_l^h \boldsymbol{\omega}_l \cdot \nabla \phi_j^x \, \mathrm{d}x + \int_K \mu_t u_l^h \phi_j^x \, \mathrm{d}x$$
$$= \int_K \mu_s w_l^h \phi_j^x \, \mathrm{d}x + \int_K f_l \phi_j^x \, \mathrm{d}x, \quad 1 \le j \le L, \forall K \in \mathcal{T}^h.$$

For (3.4.3) we use the regular finite element method. The resulting numerical method is the following: For $1 \le i \le I$,

$$\int_\Omega (w_i^h \phi_m^\omega + \alpha_i \nabla^* w_i^h \cdot \nabla^* \phi_m) \, \mathrm{d}\sigma(\boldsymbol{\omega}) = \int_\Omega u_i^h \phi_m^\omega \, \mathrm{d}\sigma(\boldsymbol{\omega}), \quad 1 \le m \le L.$$

Here $f_l(\boldsymbol{x}) := f(\boldsymbol{x}, \boldsymbol{\omega}_l)$, ν_K is the outward unit normal on ∂K, $\widehat{u_l^h}$ is the so-called numerical flux, and we may take

$$\widehat{u_l^h}(\boldsymbol{x}) = \begin{cases} u_{in}^h(\boldsymbol{x}), & \text{if } (\boldsymbol{x}, \boldsymbol{\omega}_l) \in \Gamma_- \\ \lim_{\varepsilon \to 0+} u_l^h(\boldsymbol{x} - \varepsilon \boldsymbol{\omega}_l), & \text{otherwise,} \end{cases}$$

where $u_{in}^h(\boldsymbol{x})$ is an approximation of $u_{in}(\boldsymbol{x})$ from the finite element space U^h.

More details about discretizations can be found in [11, 23].

3.5 Applications to Inverse Problems

In optical tomography, one wants to reconstruct the scattering coefficient μ_s from available measurements collected at the boundary ∂X of the domain of interest. For highly forward-peaked media, it is difficult to solve the inverse problem based upon RTE since accurate numerical solutions require a high resolution of the direction

variable. As a solution to this challenge, we consider the RT/DA as an approximation to the RTE when an inverse transport problem need to be solved:

$$\boldsymbol{\omega} \cdot \nabla u + \mu_t u = \mu_s S_j u + f. \tag{3.5.1}$$

We apply a pencil beam to a part of the boundary ∂X and get angularly resolved or averaged measurements on Γ_+. We do such experiments a few times and determine μ_s by matching the predictions from the RT/DA with the measured data on Γ_+. More precisely, let k_0 be the number of experiments. For $k = 1, \cdots, k_0$, let $\gamma_k \subset \partial X$ be the part of the spatial boundary where the pencil beam passes, let $\boldsymbol{\omega}_k \in \Omega$ be the direction of the pencil beam, and define the idealized inflow value function

$$\tilde{u}_{in,k}(\boldsymbol{x}, \boldsymbol{\omega}) = \chi_{\gamma_k}(\boldsymbol{x}) \delta(\boldsymbol{\omega} - \boldsymbol{\omega}_k). \tag{3.5.2}$$

Here, χ_{γ_k} is the characteristic function of the subset γ_k, and $\delta(\cdot)$ is the Dirac delta function. The inflow value function $u_{in,k}(\boldsymbol{x}, \boldsymbol{\omega})$ of (3.5.2) highly idealized in that the function value is innity in one direction $\boldsymbol{\omega} = \boldsymbol{\omega}_k$ and is zero along all other directions. In the inverse problems we consider here, we approximate the function $\tilde{u}_{in,k}$ by

$$u_{in,k}(\boldsymbol{x}, \boldsymbol{\omega}) = \chi_{\gamma_k}(\boldsymbol{x}) \delta_k(\boldsymbol{\omega}). \tag{3.5.3}$$

where the function $\delta_k(\boldsymbol{\omega})$ is an approximation of $\delta(\boldsymbol{\omega} - \boldsymbol{\omega}_k)$; i.e., it is nonnegative, integrates to one, and has a small support around the argument $\boldsymbol{\omega} = \boldsymbol{\omega}_k$. As an example, let $B_k \subset \Omega$ be a small neighborhood of $\boldsymbol{\omega}_k$ on Ω so that its surface area $|B_k|$ is a small positive number. Then we may take

$$\delta_k(\boldsymbol{\omega}) = \begin{cases} \frac{1}{|B_k|}, & \boldsymbol{\omega} \in B_k; \\ 0, & \boldsymbol{\omega} \in \Omega \backslash B_k. \end{cases}$$

There are two possible forms for inverse RTE problems, depending on whether the total attenuation coefficient μ_t is known or the absorption coefficient μ_a is known. In the case where μ_t is known, the forward problems are

$$\boldsymbol{\omega} \cdot \nabla u_k + \mu_t u_k = \mu_s \sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1} u_k \ \text{ in } X \times \Omega, \tag{3.5.4}$$

$$u_k = u_{in,k} \ \text{ on } \Gamma_-. \tag{3.5.5}$$

In the case where μ_a is known, we replace (3.5.4) by

$$\boldsymbol{\omega} \cdot \nabla u_k + \mu_a u_k = \mu_s \left(\sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1} u_k - u_k \right) \ \text{ in } X \times \Omega. \tag{3.5.6}$$

In this thesis, we focus our study on the case where μ_t is known. With a slight modification of the arguments, all the discussion based on (3.5.4) can be extended to that based on (3.5.6).

Let $Q_{ad} \subset Q_0 := L^2(X)$ be an admissible set for the coefficient function μ_s. Introduce the regularized functional

$$J_\varepsilon(\mu_s) = \frac{1}{2} \sum_{k=1}^{k_0} \| u_k - u_{meas,k} \|^2_{L^2_{\Gamma_+}} + \varepsilon R(\mu_s), \ \varepsilon \geq 0.$$

We assume the regularization function $R(\mu)$ satisfies a lower semicontinuity condition:

$$\mu_n \to \mu \text{ in } Q_0 \text{ as } n \to \infty \Rightarrow R(\mu) \leq \liminf_{n \to \infty} R(\mu_n). \tag{3.5.7}$$

Then the regularization inverse problem is as follows:

$$\text{Find } \mu_s \in Q_{ad} \text{ such that } J_\varepsilon(\mu_s) \text{ is minimal possible over } Q_{ad}. \tag{3.5.8}$$

There are a variety of possible choices for the regularization function $R(\mu)$, e.g., $R(\mu) = \frac{1}{2}\|u\|^2_{Q_0}$, $\frac{1}{2}\|u\|^2_{L^1(X)}$, or $\frac{1}{2}\|u\|^2_{TV}$, where $\|u\|^2_{TV}$ stands for the total varia-

tion of μ [5]. These choices satisfy the condition (3.5.7). Moreover, we may adopt the Bregman method [35, 15], , iterative regularization procedures with the use of Bregman distances, to replace the above form of (3.5.8).

We have the following result regarding problem (3.5.8).

Theorem 3.5.1. *Under the stated assumptions on the Q_{ad} and the regularization function R, for any $\varepsilon \geq 0$, problem* (3.5.8) *has a solution.*

Proof. Denote

$$m_0 := \inf_{\mu_s \in Q_{ad}} J_\varepsilon(\mu_s).$$

Let $\{\mu_n\}$ be a minimizing sequence,

$$J_\varepsilon(\mu_n) \to m_0 \quad \text{as } n \to \infty.$$

$\{\mu_n\}$ is bounded sequence in a finite dimensional space, consequently there exists a subsequence still denoting by $\{\mu_{n_m}\} \subset \{\mu_n\}$ and a function $\tilde{\mu}_s$ such that

$$\mu_{n_m} \to \tilde{\mu}_s \text{ in } L^\infty(X) \text{ as } m \to \infty, \tag{3.5.9}$$

and the limit $\tilde{\mu}_s \in Q_{ad}$ since Q_{ad} is closed.

For $1 \leq k \leq k_0$, denote by u_{k,n_m} the solution of the problems (3.5.4) with μ_s substituted by μ_{n_m}:

$$\boldsymbol{\omega} \cdot \nabla u_{k,n_m} + \mu_t u_{k,n_m} = \mu_{n_m} \sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1} u_{k,n_m} \text{ in } X \times \Omega, \tag{3.5.10}$$

$$u_{k,n_m} = u_{in,k} \quad \text{on } \Gamma_-. \tag{3.5.11}$$

Also denote by u_k the solution of the problems (3.5.4) with μ_s substituted by $\tilde{\mu}_s$:

$$\boldsymbol{\omega} \cdot \nabla u_k + \mu_t u_k = \tilde{\mu}_s \sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1} u_k \ \text{ in } X \times \Omega, \tag{3.5.12}$$

$$u_k = u_{in,k} \quad \text{on } \Gamma_-. \tag{3.5.13}$$

Then the difference $e_{k,n_m} := u_{k,n_m} - u_k$ satisfies

$$\begin{aligned} \boldsymbol{\omega} \cdot \nabla e_{k,n_m} + \mu_t u_{k,n_m} &= \mu_{n_m} \sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1} e_{k,n_m} \\ &\quad + (\mu_{n_m} - \tilde{\mu}_s) \sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1} u_k \ \text{ in } X \times \Omega, \end{aligned} \tag{3.5.14}$$

$$u_{k,n_m} = 0 \quad \text{on } \Gamma_-. \tag{3.5.15}$$

By (3.2.14), we have

$$\begin{aligned} \|e_{k,n_m}\|_{H_1^2(X\times\Omega)} &\le c \|(\mu_{n_m} - \tilde{\mu}_s) \sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1} u_k\|_{L^2(X\times\Omega)} \\ &\le c \|(\mu_{n_m} - \tilde{\mu}_s)\|_{L^2(X)} \| \sum_{i=1}^{j} \lambda_{j,i} (I - \alpha_{j,i} \Delta^*)^{-1} u_k\|_{L^2(X\times\Omega)} \\ &\to 0 \quad \text{as } m \to \infty. \end{aligned}$$

Hence for $1 \le k \le k_0$ we have

$$u_{k,n_m} \to u_k \text{ in } H_2^1(X \times \Omega) \text{ as } m \to \infty.$$

From this, we can further deduce by (1.3.1) that

$$u_{k,n_m} \to u_k \text{ in } L^2(\Gamma_+) \text{ as } m \to \infty.$$

Now consider $J_\varepsilon(\tilde{\mu}_n)$:

$$\begin{aligned}
J_\varepsilon(\tilde{\mu}_s) &= \frac{1}{2}\sum_{k=1}^{k_0} \|u_k - u_{meas,k}\|^2_{L^2(\Gamma_+)} + \varepsilon R(\tilde{\mu}_s) \\
&= \frac{1}{2}\sum_{k=1}^{k_0} \|(u_{k,n_m} - u_{meas,k}) + (u_k - u_{k,n_m})\|^2_{L^2(\Gamma_+)} + \varepsilon R(\tilde{\mu}_s) \\
&= \frac{1}{2}\sum_{k=1}^{k_0} \|(u_{k,n_m} - u_{meas,k})\|^2_{L^2(\Gamma_+)} + \varepsilon R(\mu_{n_m}) \\
&\quad + \sum_{k=1}^{k_0} (u_{k,n_m} - u_{meas,k}, u_k - u_{k,n_m})_{L^2(\Gamma_+)} + \frac{1}{2}\sum_{k=1}^{k_0} \|(u_k - u_{k,n_m})\|^2_{L^2(\Gamma_+)} \\
&\quad + \varepsilon(R(\tilde{\mu}_s) - R(\mu_{n_m})) \\
&= J_\varepsilon(\mu_{n_m}) \\
&\quad + \sum_{k=1}^{k_0} (u_{k,n_m} - u_{meas,k}, u_k - u_{k,n_m})_{L^2(\Gamma_+)} + \frac{1}{2}\sum_{k=1}^{k_0} \|(u_k - u_{k,n_m})\|^2_{L^2(\Gamma_+)} \\
&\quad + \varepsilon(R(\tilde{\mu}_s) - R(\mu_{n_m})).
\end{aligned}$$

Since

$$(u_{k,n_m} - u_{meas,k}, u_k - u_{k,n_m})_{L^2(\Gamma_+)} \le \|u_{k,n_m} - u_{meas,k}\|^2_{L^2(\Gamma_+)} \|u_k - u_{k,n_m}\|^2_{L^2(\Gamma_+)},$$

so we have

$$\begin{aligned}
J_\varepsilon(\tilde{\mu}_s) \le J_\varepsilon(\mu_{n_m}) &\\
&+ \sum_{k=1}^{k_0} \|u_{k,n_m} - u_{meas,k}\|^2_{L^2(\Gamma_+)} \|u_k - u_{k,n_m}\|^2_{L^2(\Gamma_+)} \\
&+ \frac{1}{2}\sum_{k=1}^{k_0} \|(u_k - u_{k,n_m})\|^2_{L^2(\Gamma_+)} + \varepsilon(R(\tilde{\mu}_s) - R(\mu_{n_m})).
\end{aligned}$$

From (3.5.7), we know $R(\tilde{\mu}_s) - R(\mu_{n_m}) \le 0$. Let $m \to \infty$ in the above inequality, we have

$$J_\varepsilon(\tilde{\mu}_s) \le \liminf_{j\to\infty} J_\varepsilon(\mu_{n_m}).$$

Therefore, $J_\varepsilon(\tilde{\mu}_s) = m_0$, and $\tilde{\mu}_s$ is the solution of (3.5.8). □

3.6 Numerical Examples

We present some preliminary numerical examples on solving RT/DA and inverse RTE problems using RT/DA as the forward model with angularly resolved measurement. The purpose is to show that $\mathrm{RT/DA}_j$ is a good approximation to the RTE, and to demonstrate that the method described in previous sections is suitable for reconstruction of the scattering coefficient.

3.6.1 Forward Problems

First we consider solving a problem of type (3.3.3):

$$w(\boldsymbol{\omega}) - \alpha\Delta^* w(\boldsymbol{\omega}) = f(\boldsymbol{\omega}),\ \boldsymbol{\omega} \in \Omega, \tag{3.6.1}$$

where $\alpha \geq 0$ and $f \in L^2(\Omega)$ are given. The corresponding weak formulation is

$$w \in H^1(\Omega):\ \int_\Omega (wv + \alpha\nabla^* w \cdot \nabla^* v)\,\mathrm{d}\sigma(\boldsymbol{\omega}) = \int_\Omega fv\,\mathrm{d}\sigma(\boldsymbol{\omega})\ \forall v \in H^1(\Omega). \tag{3.6.2}$$

We use the standard spherical coordinates, $x(\psi,\theta) = \cos\psi\sin\theta$, $y(\psi,\theta) = \sin\psi\sin\theta$, $z(\psi,\theta) = \cos\theta$ for $0 \leq \psi \leq 2\pi$, $0 \leq \theta \leq \pi$, and follow [2] to formulate a Galerkin method where the basis functions are piecewise linear on triangular elements of the mesh and bilinear on rectangular elements of the mesh. The partition of the closed domain $G = [0, 2\pi] \times [0, \pi]$ for the variables ψ and θ in [2] involves two parameters: n_ϕ and n_θ, where n_ϕ is usually taken to be 4, 6 or 8 and n_θ is even. Initially, the domain G is divided n_ϕ times in the ψ-direction and n_θ times in the θ-direction. The generated rectangles at $\theta = 0, \pi$ remain unchanged, corresponding to the triangles on the unit sphere at the poles. Define $h_\theta = \pi/n_\theta$. An edge with $\theta = kh_\theta \leq \pi/2$, $1 \leq k \leq n_\theta/2$, is equally split into $(k-1)$ parts. Then the generated nodes are connected properly, and

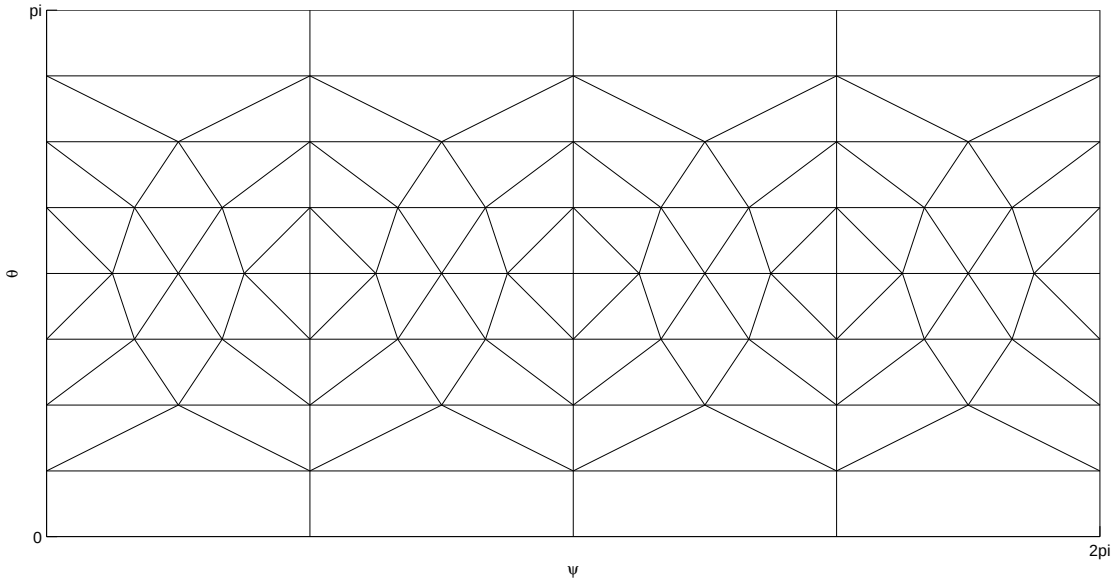

Figure 3.5. Angular mesh with 66 nodes.

the generated mesh is reflected with respect to the line $\theta = \pi/2$. The corresponding mesh on the unit sphere Ω has $n_\theta^2 n_\phi/2$ elements and $2 + n_\theta^2 n_\phi/4$ nodes. For reference, the angular mesh using $n_\phi = 4$ and $n_\theta = 8$ with 66 nodes is shown in Fig. 3.5 and 3.6. We refer the reader to [2] for details of this method.

In the following example, we compare the numerical solution of the RTE with the numerical solutions to the RT/DA$_j$ equation calculated on the same mesh. This leaves us with a choice of weights when solving the RTE. Initially, the choice was made that $w_i = \frac{4\pi}{N}$ where N is the number of angular nodes. However, this is not a good quadrature rule, as the nodes are not quite evenly spaced on the sphere. When solving the approximation to the RT/DA$_j$ equation, a matrix A is formed with the

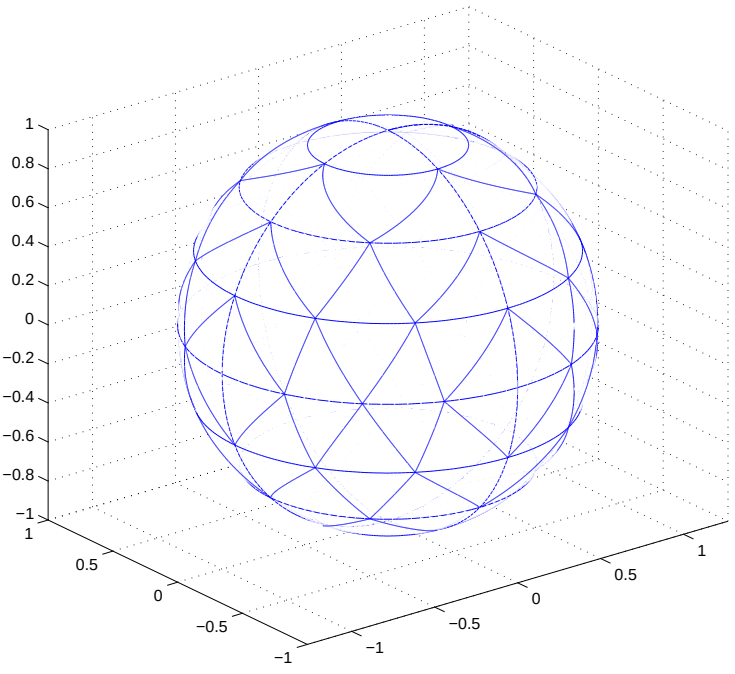

Figure 3.6. Angular mesh with 66 nodes.

property that

$$f^T A g = \int_\Omega f(\boldsymbol{\omega}) g(\boldsymbol{\omega}) \, \mathrm{d}\sigma(\boldsymbol{\omega})$$

if f is a vector containing function values of f at the nodes $\boldsymbol{\omega}_i$, and f, g are elements of the finite element space associated with the angular mesh. We choose the weight vector w to be $w = Ae$, where e is the vector with all components 1. This quadrature rule is exact for all functions in the finite element space associated with the angular mesh. Since this rule correctly integrates any piecewise linear function in the finite element space, it may be thought of as an analogue of the trapezoidal rule for the sphere.

For a positive integer n, we partition X into n^3 subcubes $\{X_i\}$, each with edge length $1/n$. Denote by S the set of all the centers and vertices of the subcubes. A mesh is generated by creating the Delaunay tessellation of the points in S. This is accomplished using the delaunay3 algorithm in MATLAB. The option QJ is passed to delaunay3 to ensure that no hanging nodes are created. Denote by h the maximum length of the edges of the tetrahedron in the mesh; $h = \sqrt{2}/n$. The local polynomial degree $k = 1, 2,$ or 3 in different examples. In all the following cases, we will investigate the error

$$e := \left\{ \frac{4\pi}{L} \sum_l \| u_h^l - u(\cdot, \boldsymbol{\omega}_l) \|_{L^2(X)}^2 \right\}^{\frac{1}{2}}.$$

Example 3.6.1. Since we have

$$S_2 \widetilde{Y}_2^1 = g^2 \widetilde{Y}_2^1 \text{ and } S\widetilde{Y}_2^1 = g^2 \widetilde{Y}_2^1$$

where

$$S_2 = \lambda_1(I - \alpha_1\Delta^*)^{-1} + \lambda_2(I - \alpha_2\Delta^*)^{-1},$$

we choose the function f as

$$f = 10\Big(\pi\big(\omega_1\cos(\pi x_1)\sin(\pi x_2)\sin(\pi x_3) + \omega_2\sin(\pi x_1)\cos(\pi x_2)\sin(\pi x_3) + \omega_3\sin(\pi x_1)\sin(\pi x_2)\cos(\pi x_3)\big) + (1-g^2)\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)\Big) \sin(\theta)\cos(\theta)\cos(\varphi),$$

so that the true solution u to two terms $\mathrm{RT}/\mathrm{DA}_2$ is

$$\begin{aligned} u &= 10\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)\widetilde{Y}_2^1(\theta,\varphi) \\ &= 10\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)\sin(\theta)\cos(\theta)\cos(\varphi). \end{aligned}$$

Here

$$\widetilde{Y}_2^1(\theta,\varphi) = \sin(\theta)\cos(\theta)\cos(\varphi)$$

is the spherical harmonic function of order 2 **without** normalization. Observe that u is also the solution to the RTE with the same f and u_{in}. We have $\|u\|_{L^2} \approx 3.23$. Take $g = 0.95$. Numerical results are reported in Table 3.1, Table 3.2, and Table 3.3.

Example 3.6.2. We choose the function f as

$$f = 10\Big(\pi\big(\omega_1\cos(\pi x_1)\sin(\pi x_2)\sin(\pi x_3) + \omega_2\sin(\pi x_1)\cos(\pi x_2)\sin(\pi x_3) + \omega_3\sin(\pi x_1)\sin(\pi x_2)\cos(\pi x_3)\big) + (1-g^3)\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)\Big) \sin(\theta)(5\cos^2(\theta)-1)\cos(\phi),$$

h	one term	two terms	three terms
$\frac{\sqrt{2}}{8}$	0.087978897443878	0.030754566550038	0.029800374933836
$\frac{\sqrt{2}}{10}$	0.084868360282745	0.022429670313491	0.021072765036380
$\frac{\sqrt{2}}{12}$	0.083645955674038	0.018452962498677	0.016765264925806
$\frac{\sqrt{2}}{14}$	0.083085608812000	0.016421820733581	0.014494170738487

Table 3.1. L^2 error for several values of h with $n_\phi = 4$ and $n_\theta = 16$, linear basis

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.093479394610576	0.048405700382365	0.048354199319172
$\frac{\sqrt{2}}{8}$	0.083278744815350	0.028099566473783	0.027929407099540
$\frac{\sqrt{2}}{10}$	0.080020666881298	0.018751082946316	0.018463623001191
$\frac{\sqrt{2}}{12}$	0.078738745460190	0.013822954117419	0.013415984756831

Table 3.2. L^2 error for several values of h with $n_\phi = 8$ and $n_\theta = 16$, linear basis

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.077413406135051	0.006623005104947	0.005701175153446
$\frac{\sqrt{2}}{8}$	0.077369214745947	0.006201823846530	0.005204040597438
$\frac{\sqrt{2}}{10}$	0.077361619203407	0.006132124349994	0.005120507073922
$\frac{\sqrt{2}}{12}$	0.077359698106804	0.006115578856959	0.005100624018312

Table 3.3. L^2 error for several values of h with $n_\phi = 8$ and $n_\theta = 16$, quadratic basis

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.463734689812156	0.127073397266084	0.120410563943650
$\frac{\sqrt{2}}{8}$	0.449646262849763	0.081898815212195	0.072279479250396
$\frac{\sqrt{2}}{10}$	0.445209464442630	0.063384272629793	0.050981264861603
$\frac{\sqrt{2}}{12}$	0.443433547889453	0.054971894574604	0.040416990016028
$\frac{\sqrt{2}}{14}$	0.442599322952157	0.050840372347826	0.034816783868873

Table 3.4. L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.441464298779398	0.045681099045607	0.027228664779152
$\frac{\sqrt{2}}{8}$	0.441405557517200	0.045313371791585	0.026624500026725
$\frac{\sqrt{2}}{10}$	0.441394862234064	0.045253496801470	0.026526455429515
$\frac{\sqrt{2}}{12}$	0.441392048516620	0.045239147898322	0.026503177084603

Table 3.5. L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$, quadratic basis

so that the true solution u to three terms $\mathrm{RT}/\mathrm{DA}_3$ is

$$\begin{aligned} u &= 10\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)\widetilde{Y}_3^1(\theta,\varphi) \\ &= 10\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)\sin(\theta)(5\cos^2(\theta)-1)\cos(\phi). \end{aligned}$$

Here

$$\widetilde{Y}_3^1(\theta,\varphi) = \sin(\theta)(5\cos^2(\theta)-1)\cos(\phi)$$

is the spherical harmonic function of order 3 **without** normalization. Observe that u is also the solution to the RTE with the same f and u_{in}. We have $\|u\|_{L^2} \approx 7.73$. Take $g = 0.95$. Numerical results are reported in Table 3.4, Table 3.5, Table 3.6 and Table 3.7.

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.441390514480059	0.045232710729941	0.026492985253896
$\frac{\sqrt{2}}{8}$	0.441390383151706	0.045231839309043	0.026491542098902
$\frac{\sqrt{2}}{10}$	0.441390368460797	0.045231757646604	0.026491410891150
$\frac{\sqrt{2}}{12}$	0.441390366605574	0.045231745448037	0.026491391281795

Table 3.6. L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 12$, cubic basis

(n_ϕ, n_θ)	one term	two terms	three terms
$(10, 20)$	0.427826870050316	0.031948671006974	0.015003705436903
$(10, 24)$	0.423046760764947	0.028406067302392	0.012158348859085
$(10, 28)$	0.420028091697620	0.026192199368362	0.010355256884040
$(10, 32)$	0.417984558985139	0.024730199840922	0.009165684779992
$(10, 34)$	0.417199441844181	0.024182081021210	0.008726277783270
$(10, 36)$	0.416529355185597	0.023720564520087	0.008359595839723
$(10, 38)$	0.415952773797562	0.023330771535111	0.008057419677351
$(10, 40)$	0.415452261677357	0.022996709810009	0.007802255310022

Table 3.7. L^2 error from different values of n_θ with $h = \frac{\sqrt{2}}{6}$, quadratic elements

Comparing the results from Table 3.4, Table 3.5 and Table 3.6, notice that using quadratic elements for spatial variable even relatively small values of $h = 6$ can achieve a satisfactory accuracy of spatial resolution. So, as shown in Table 3.7, we compare the results for different values of n_θ while fixing $h = \frac{\sqrt{2}}{6}$. Observe as the value of n_θ is increased, the ratio of errors between two terms and three terms is also increased.

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.513412329172713	0.124304890455056	0.091099261007513
$\frac{\sqrt{2}}{8}$	0.506379422514688	0.103081753950269	0.061596139910172
$\frac{\sqrt{2}}{10}$	0.504115567615279	0.096002570299699	0.050058219142452
$\frac{\sqrt{2}}{12}$	0.503188663209177	0.093182269602377	0.045029981152275
$\frac{\sqrt{2}}{14}$	0.502744359514141	0.091889468995365	0.042629778599811

Table 3.8. L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 12$, linear basis

Example 3.6.3. We choose the function f as

$$f = 10\Big(\pi\big(\omega_1\cos(\pi x_1)\sin(\pi x_2)\sin(\pi x_3) + \omega_2\sin(\pi x_1)\cos(\pi x_2)\sin(\pi x_3) + \omega_3\sin(\pi x_1)\sin(\pi x_2)\cos(\pi x_3)\big) + (1-g^4)\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)\Big) \sin(\theta)(7\cos^3(\theta) - 3\cos(\theta))\cos(\phi),$$

so that the true solution u to four terms RT/DA$_4$ is

$$\begin{aligned} u &= 10\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)\widetilde{Y}_4^1(\theta,\varphi) \\ &= 10\sin(\pi x_1)\sin(\pi x_2)\sin(\pi x_3)\sin(\theta)(7\cos^3(\theta) - 3\cos(\theta))\cos(\phi). \end{aligned}$$

Here

$$\widetilde{Y}_4^1(\theta,\varphi) = \sin(\theta)(7\cos^3(\theta) - 3\cos(\theta))\cos(\phi)$$

is the spherical harmonic function of order 4 **without** normalization. Observe that u is also the solution to the RTE with the same f and u_{in}. We have $\|u\|_{L^2} \approx 5.28$. Choose $g = 0.95$. Numerical results are reported in Table 3.8, Table 3.9, Table 3.10, Table 3.12, and Table 3.13.

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.502108767307019	0.090297881912054	0.039697916305660
$\frac{\sqrt{2}}{8}$	0.502078777536988	0.090200055886271	0.039497558630619
$\frac{\sqrt{2}}{10}$	0.502073208545302	0.090183630747193	0.039464879908366
$\frac{\sqrt{2}}{12}$	0.502071702890121	0.090179564477488	0.039457031268414

Table 3.9. L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 12$, quadratic basis

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.488801523121195	0.111232479155214	0.085577264227339
$\frac{\sqrt{2}}{8}$	0.481543506929978	0.087453082460270	0.053837237563101
$\frac{\sqrt{2}}{10}$	0.479211792853763	0.079210903923337	0.040468179675268
$\frac{\sqrt{2}}{12}$	0.478259006361768	0.075871893105123	0.034227149115468

Table 3.10. L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$, linear basis

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.477154270077980	0.072448143198620	0.027125273078275
$\frac{\sqrt{2}}{8}$	0.477123504685512	0.072329454571122	0.026837624433798
$\frac{\sqrt{2}}{10}$	0.477117710492082	0.072309566669271	0.026790849841183
$\frac{\sqrt{2}}{12}$	0.477116157625644	0.072304676150827	0.026779678153902

Table 3.11. L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$, quadratic basis

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.497048425462586	0.117659769002802	0.089181098367369
$\frac{\sqrt{2}}{8}$	0.489935688201531	0.095563258349795	0.059555381370790
$\frac{\sqrt{2}}{10}$	0.487655654311102	0.088126084493897	0.047911541878160
$\frac{\sqrt{2}}{12}$	0.486725034795299	0.085159144888163	0.042823240293432

Table 3.12. L^2 error for several values of h with $n_\phi = 4$ and $n_\theta = 16$, linear basis

h	one term	two terms	three terms
$\frac{\sqrt{2}}{6}$	0.483655065676352	0.110015215553379	0.085959538256459
$\frac{\sqrt{2}}{8}$	0.476386096483540	0.086204145231700	0.054883508602356
$\frac{\sqrt{2}}{10}$	0.474056908664568	0.077968381850791	0.042086706857876
$\frac{\sqrt{2}}{12}$	0.473107002464645	0.074641256847616	0.036253595211279

Table 3.13. L^2 error for several values of h with $n_\phi = 4$ and $n_\theta = 20$, linear basis

3.6.2 Forward Problems Using Optimized Parameters

In this section we repeat some examples in the previous subsection to the impact of optimized parameters α_i and λ_i.

Example 3.6.4. In this example, the geometry settings are the same as in Example 3.6.1, while we use the optimized parameters α_i and λ_i for three terms approximation introduced in section before. Numerical results are reported in Table 3.14. Compared to the results of three terms approximation in Table 3.3, we can see that the accuracy of the three terms approximation is improved using optimized parameters.

Example 3.6.5. In this example, the geometry settings are the same as in Example 3.6.3, while we use the optimized parameters α_i and λ_i for three terms approximation

h	three terms
$\frac{\sqrt{2}}{6}$	0.005472555014551
$\frac{\sqrt{2}}{8}$	0.004951879883814
$\frac{\sqrt{2}}{10}$	0.004863916725154
$\frac{\sqrt{2}}{12}$	0.004842957090746

Table 3.14. L^2 error for several values of h with $n_\phi = 8$ and $n_\theta = 16$, quadratic basis

h	three terms
$\frac{\sqrt{2}}{6}$	0.033996539867789
$\frac{\sqrt{2}}{8}$	0.033765013224770
$\frac{\sqrt{2}}{10}$	0.033727357895215
$\frac{\sqrt{2}}{12}$	0.033718347163895

Table 3.15. L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 12$, quadratic basis

introduced in section before. Numerical results are reported in Table 3.15 and Table 3.16. Compared to the results of three terms approximation in Table 3.9 and Table 3.11, we can see that the accuracy of the three terms approximation is improved using optimized parameters.

3.6.3 Inverse Problems

Example 3.6.6. Consider the inverse RTE problem using RT/DA of recovering μ_s. $u_{meas,k}$ is an approximate solution obtained by solving the RTE with the following parameters:

$$\mu_s = \begin{cases} 1.1 & \text{if } \boldsymbol{x} \in R_1, \\ 1.0 & \text{otherwise,} \end{cases} \qquad \mu_t = \mu_s + \mu_a = \begin{cases} 1.2 & \text{if } \boldsymbol{x} \in R_1, \\ 1.2 & \text{otherwise,} \end{cases} \tag{3.6.3}$$

h	three terms
$\frac{\sqrt{2}}{6}$	0.022779826642180
$\frac{\sqrt{2}}{8}$	0.022439974593549
$\frac{\sqrt{2}}{10}$	0.022384793430021
$\frac{\sqrt{2}}{12}$	0.022371655820664

Table 3.16. L^2 error for several values of h with $n_\phi = 6$ and $n_\theta = 16$, quadratic basis

$$u_{meas,k}(\boldsymbol{x},\boldsymbol{\omega})|_{\Gamma_-} = \begin{cases} \frac{L}{4\pi} & \text{if } \boldsymbol{x} \in S_k \text{ and } \boldsymbol{\omega} \text{ is around } \boldsymbol{\omega}_k, \\ 0 & \text{otherwise,} \end{cases} \quad k = 1, 2, \cdots, 6.$$

The domain $X = (-1, 1)^3$. The domains $R_1 = X \cap \{(x, y, z)|x + y < 0\}$, $R_2 := X \backslash R_1$ and the surfaces $S_1, S_2, \cdots, S_6$ are illustrated in Figure 3.7. We use Henyey-Greenstein function as the phase function with anisotropy factor $g = 0.9$. The spatial mesh used generate $u_{meas,k}$ is shown in each of the preceding figures and has 768 elements. The angular mesh has $n_\theta^2 n_\phi/2$ elements and $2 + n_\theta^2 n_\phi/4$ nodes. Since the anisotropy factor g is relatively closed to 1, we set $\phi = 6$ and $\theta = 50$. Then there are 3752 nodes on the unit sphere and hence in the above definition $L = 3752$. Directions $\boldsymbol{\omega}_i, i = 1, 2, \cdots, 6$ are node points of the angular mesh that point from the centers of $S_i, i = 1, 2, \cdots, 6$, respectively, to the center of the cube X.

In order to reconstruct μ_s, we use approximate solutions to the RT/DA$_j$ constructed on a spatial mesh with 96 elements which is shown in Figure 3.8. Define

$$Q_{ad} = \{\mu \mid \mu|_{R_k} = \mu_k \in \mathbb{R}, 0 \le \mu_k \le \mu_t, k = 1, 2\}.$$

In all cases, local polynomials of order 2 for spacial basis are used. Since six sets of measurements are used to determine two constants that define μ_s, we may take

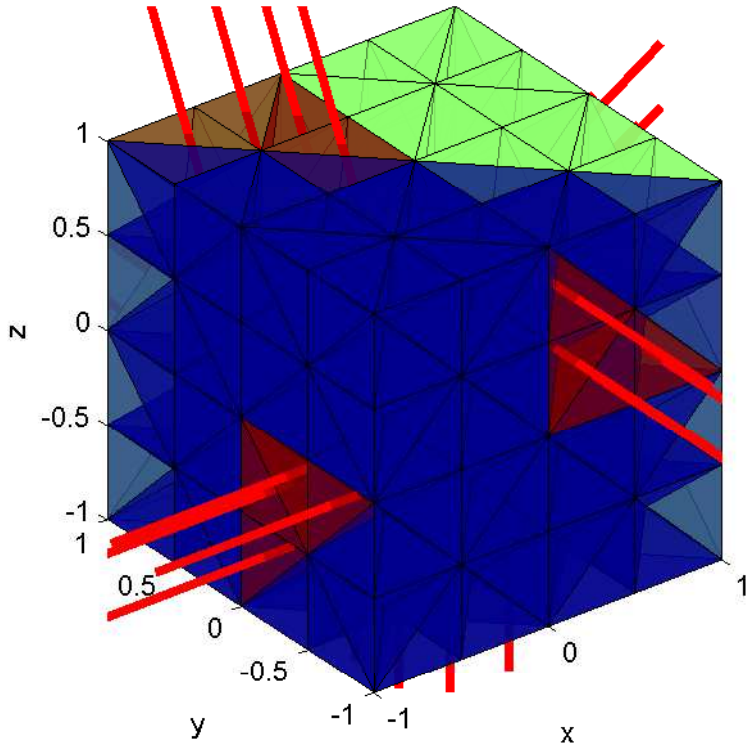

Figure 3.7. Domains R_1 and R_2 in different colors and pencil beams.

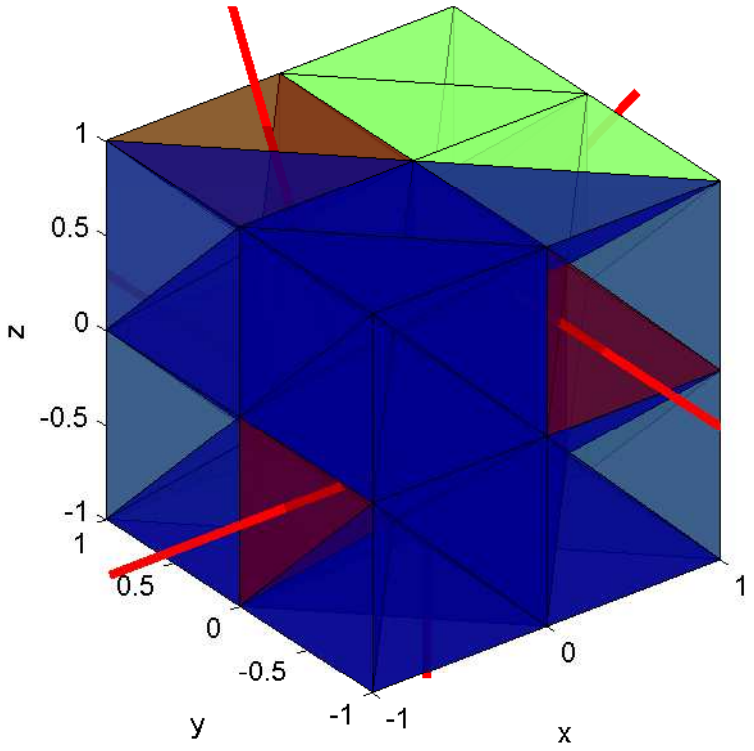

Figure 3.8. Domains R_1 and R_2 in different colors and pencil beams.

$\varepsilon = 0$. To study the effect of noise to the reconstruction, certain amount of random error to $u_{meas,k}$ was deliberately added in some cases. 'U% noise' means letting err be an uniformly distributed random variable supported on $[-U/100, U/100]$, we use $(1 + err)u_{meas,k}$ as the new $u_{meas,k}$ for reconstruction. Denote the reconstruction of μ_s by μ_s^r. Denote the reconstruction of μ_s by μ_s^r. The numerical results are reported in Table 3.17–3.22.

j	one term	two terms	three terms
$\mu_{s,1}^r$	1.307332148117757	1.184649162111553	1.140950919769447
$\mu_{s,2}^r$	1.130559023576054	1.028977417310920	0.995198063271455

Table 3.17. Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, no noise

j	one term	two terms	three terms
$\mu_{s,1}^r$	1.307445398513650	1.184877193570524	1.140972655098905
$\mu_{s,2}^r$	1.132561461069689	1.032481669554029	0.999970577474685

Table 3.18. Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 5% noise

j	one term	two terms	three terms
$\mu_{s,1}^r$	1.307553712970200	1.184537149772306	1.140381351606754
$\mu_{s,2}^r$	1.134547150148219	1.037527801783728	1.006477991366319

Table 3.19. Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 10% noise

j	one term	two terms	three terms
$\mu_{s,1}^r$	1.307987577940499	1.185016274645279	1.141569524983130
$\mu_{s,2}^r$	1.142351989092040	1.051924403202183	1.022939817387053

Table 3.20. Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 30% noise

j	one term	two terms	three terms
$\mu_{s,1}^r$	1.308526001461540	1.185511840058246	1.142128899462032
$\mu_{s,2}^r$	1.150037596691839	1.065899499792228	1.041020273065097

Table 3.21. Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 50% noise

j	1	2	3
$\mu_{s,1}^r$	1.309250935645982	1.186029515462364	1.142543802847374
$\mu_{s,2}^r$	1.157762819796739	1.079462195447448	1.058651576122694

Table 3.22. Reconstructed values of μ_s, with $g = 0.9$, $\mu_t = 1.2$, 70% noise

Example 3.6.7. The geometric setting is the same as in the previous example. However in this example, we assume that values of μ_a rather than of μ_t are known. The functions $u_{meas,k}$ are calculated on the same domain with same spacial and angular mesh using the data

$$\mu_s = \begin{cases} 1.1 & \text{if } \boldsymbol{x} \in R_1, \\ 1.0 & \text{otherwise,} \end{cases} \qquad \mu_a = 0.01,$$

$$\mu_t = \mu_s + \mu_a = \begin{cases} 1.11 & \text{if } \boldsymbol{x} \in R_1, \\ 1.01 & \text{otherwise,} \end{cases}$$

The numerical results are reported in Table 3.23–3.25.

j	1	2	3
$\mu_{s,1}^r$	1.037578589049473	1.077148233073228	1.102254956113231
$\mu_{s,2}^r$	0.960043544323594	0.973029333399142	0.981394330387275

Table 3.23. Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, no noise

j	1	2	3
$\mu_{s,1}^r$	1.038772793632809	1.078296520273898	1.103380478276057
$\mu_{s,2}^r$	0.944865530405467	0.957440809505007	0.965746953952433

Table 3.24. Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 10% noise

j	1	2	3
$\mu_{s,1}^r$	1.040974797806643	1.080511406866368	1.105460861887335
$\mu_{s,2}^r$	0.915524554010455	0.927264000241060	0.935158011227946

Table 3.25. Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 30% noise

Example 3.6.8. In this example, we consider the inverse RTE problem using RT/DA of recovering three different values of μ_s in three regions assuming μ_a is known. The geometric setting is the same as in Example 3.6.6. $u_{meas,k}$ is an approximate solution obtained by solving the RTE with the following parameters:

$$\mu_s = \begin{cases} 1.1 & \text{if } \boldsymbol{x} \in R_1, \\ 1.0 & \text{if } \boldsymbol{x} \in R_2, \\ 0.9 & \text{otherwise,} \end{cases}, \qquad \mu_a = 0.01, \tag{3.6.4}$$

$$\mu_t = \mu_s + \mu_a = \begin{cases} 1.11 & \text{if } \boldsymbol{x} \in R_1, \\ 1.01 & \text{if } \boldsymbol{x} \in R_2, \\ 0.91 & \text{otherwise,} \end{cases} \tag{3.6.5}$$

$$u_{meas,k}(\boldsymbol{x}, \boldsymbol{\omega})|_{\Gamma_-} = \begin{cases} \frac{L}{4\pi} & \text{if } \boldsymbol{x} \in S_k \text{ and } \boldsymbol{\omega} \text{ is around } \boldsymbol{\omega}_k, \\ 0 & \text{otherwise,} \end{cases} \quad k = 1, 2, \cdots, 5.$$

The domain $X = (-1, 1)^3$. The domains $R_1 = X \cap \{(x, y, z) | x + y < 0\}$, $R_2 := X \cap \{(x, y, z) | x + y > 0 \text{ and } x - y < 0\}$, $R_3 = X \backslash (R_1 \cup R_2)$ and the surfaces $S_1, S_2, \cdots, S_6$ are illustrated in Figure 3.9. The spatial and angular mesh used to generate $u_{meas,k}$ are also the same as example 3.6.6. Denote the reconstruction of μ_s by μ_s^r. The numerical results are reported in Table 3.26–3.28.

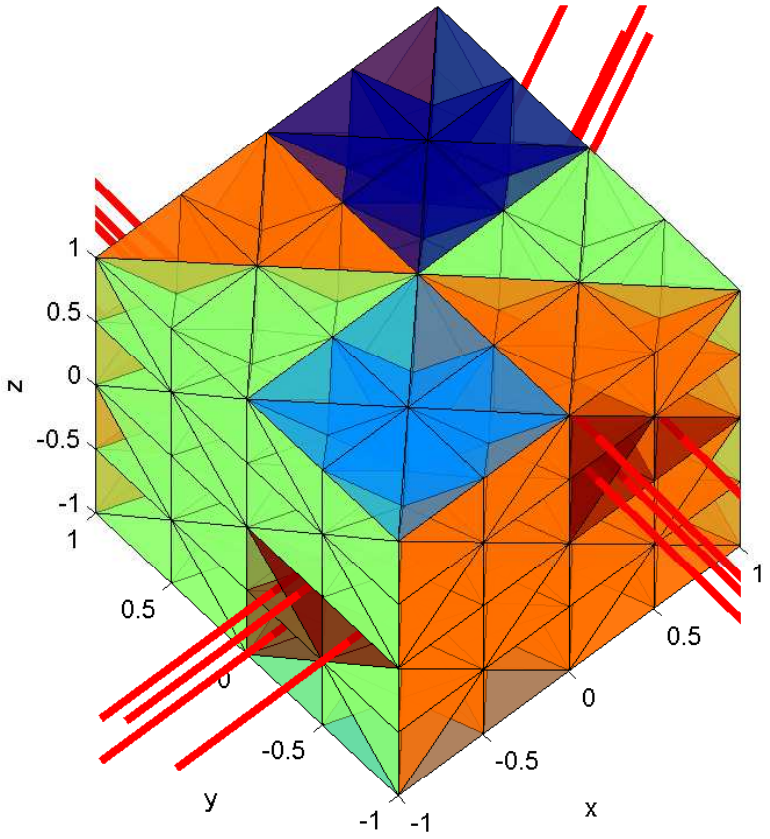

Figure 3.9. Domains R_1, R_2 and R_3 in different colors and pencil beams.

j	1	2	3
$\mu_{s,1}^r$	1.042982053786405	1.116008455044000	1.162839453859029
$\mu_{s,2}^r$	0.969419245373666	0.993990755662393	1.009987504246735
$\mu_{s,3}^r$	0.839130505169611	0.861313671548887	0.875408053010619

Table 3.26. Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, no noise

j	1	2	3
$\mu_{s,1}^r$	0.996045888833225	1.152158224842208	1.199220570048405
$\mu_{s,2}^r$	0.964991021328478	0.986210184892645	1.002006897620372
$\mu_{s,3}^r$	0.830484981851958	0.850240473639506	0.864240578253314

Table 3.27. Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 20% noise

j	1	2	3
$\mu_{s,1}^r$	1.014845135401424	1.188191269975630	1.236466996750117
$\mu_{s,2}^r$	0.949268582503551	0.979009743361311	0.994556983400018
$\mu_{s,3}^r$	0.846131104063377	0.839058589171677	0.852780805908377

Table 3.28. Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 40% noise

Example 3.6.9. All the settings are the same as previous example except we use 6 pencil beams and take 6 different measurement $u_{meas,k}|_{\Gamma_-}, k = 1, 2, \cdots, 6$ instead of 5 measurement in the previous example. The numerical results are reported in Table 3.29–3.31. We see improvements as the number of measurements increases.

j	1	2	3
$\mu_{s,1}^r$	1.005858890571891	1.078716424887745	1.124124971967051
$\mu_{s,2}^r$	0.969053207228378	0.994972368450655	1.011642109690253
$\mu_{s,3}^r$	0.837301482387151	0.860444738446235	0.875353737896193

Table 3.29. Reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, no noise

j	1	2	3
$\mu_{s,1}^r$	1.019652738209183	1.093223998277784	1.138933695096912
$\mu_{s,2}^r$	0.965090775223684	0.990935141644198	1.007538947099780
$\mu_{s,3}^r$	0.832674306332647	0.855720258267141	0.870572175497679

Table 3.30. True and reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 10% noise

j	1	2	3
$\mu_{s,1}^r$	1.033458950012950	1.107933118962366	1.153659105414961
$\mu_{s,2}^r$	0.961337791422376	0.986917309559296	1.003526676943794
$\mu_{s,3}^r$	0.828152557217201	0.850860483927568	0.865503158058620

Table 3.31. True and reconstructed values of μ_s, with $g = 0.9$, $\mu_a = 0.01$, 20% noise

CHAPTER 4

CONCLUSIONS AND FURTHER WORK

In this thesis we study the approximations to Radiative Transfer Equations. For the Fokker-Planck equation, the well-posedness of the problem is shown. A family of differential approximations of the RTE is then discussed. Existence and uniqueness of the problem are proved and a numerical method is given. Finally their applications to inverse problems are studied. For both forward and inverse problems numerical results are given to demonstrate the effectiveness of the approximations.

There are several topics that warrant further investigation related to this work. Even though some convergence analysis was performed in the paper [22], a complete error analysis is still missing. Specifically, I plan to study the error analysis of finite element method for the equation $(I - \alpha_i \Delta^*)w = u$. The finite element method is described in [2]. Since Δ^* is the Laplace-Beltrami operator on the unit sphere, its error analysis is quite different from one in the Euclidean space. A posteriori error estimate is almost done and the priori one is worth effort for research.

While it is reasonable to expect uniqueness when the number of measurements is equal to or larger than the number of unknowns of the scattering coefficient, rigorous theoretical results remain to be done. These results are also helpful for practical reconstruction of the scattering parameter since it is desirable to use the least number of measurements that guarantees uniqueness. In [18], no regularization is used, since the number of measurements is equal to the number of unknowns needed to determine

the scattering parameter. In more realistic simulations, the number of measurements is much smaller than the number of unknowns for the scattering parameter, and hence regularization plays a fundamental role in dealing with the ill-conditioning of the problem and in controlling data noise. While it is believed that the Bregman-type methods [35, 15] are prominent choices for the type of inverse problems, detailed theoretical analysis and simulation experiments are needed to identify the best choice of the regularization function.

In biomedical imaging, traditionally, more data provides more information, and reconstruct more accurate images. Therefore one has to collect a lot of data from measurements. They are however very expensive to acquire in many medical imaging modalities. Many new approaches have been developed to make better use of available data. One of them is the so-called sparse recovery. Sparsity has gained a lot of interest in recent years. It can be exploited to capture or recover information efficiently by using only a small number of measurements. I am planning to apply the sparse recovery method to BI problems. I will implement algorithms of sparse recovery to BI and will study the theoretical aspects.

Using the iteration method of the differential approximations to radiative transfer equations we introduced in this thesis, the computation on the spatial variable $\boldsymbol{x}$ and angular variable $\boldsymbol{\omega}$ are decoupled and the numerical solution of the angular variable can be solved simultaneously on all the spatial nodes. Therefore this method can be implemented with very high parallel efficiency on the scale of thousands of process units. These properties make the so-called GPU computing a perfect candidate

to implement the method presented in this thesis.

REFERENCES

[1] A. Agoshkov. *Boundary Value Problems for Transport Equations*. Birkhäuser, Boston, 1998.

[2] T. APEL and C. PESTER. Clement-type interpolation on spherical domains—interpolation error estimates and application to a posteriori error estimation. *IMA J. Numer. Anal.*, 25:310–336, 2005.

[3] K. Atkinson and W. Han. *Theoretical Numerical Analysis: A Functional Analysis Framework*. Springer, New York, 3rd edition, 2009.

[4] K. Atkinson and W. Han. *Spherical Harmonics and Approximations on the Unit Sphere: An Introduction*, volume 2044 of *Lecture Notes in Mathematics*. Springer-Verlag, 2012.

[5] H. Attouch, G. Buttazzo, and G. Michaille. *Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization*. MOS-SIAM Series on Optimization. SIAM, Philadelphia, 2006.

[6] G. Avtandilov, A. Dembo, O. Komardin, P. Lazarev, M. Paukshto, L. Shkolnik, and O. Zayratiyants. Human tissue analysis by small-angle X-ray scattering. *Journal of Applied Crystallography*, 33:511–513, 2000.

[7] G. Bal and A. Tamasan. Inverse source problems in transport equations. *SIAM J. Math. Anal.*, 39:57–76, 2007.

[8] M. Caro and J. Ligou. Treatment of scattering anisotropy of neutrons through the Boltzmann-Fokker-Planck equation. *Nucl. Sci. Eng.*, 83:242–250, 1983.

[9] D. Chapman, W. Thomlinson, R. E. Johnston, D. Washburn E. Pisano, N. Gmür, Z. Zhong, R. Menk, F. Arfelli, and D. Sayers. Diffraction enhanced X-ray imaging. *Physics in Medicine and Biology*, 42:2015–2025, 1997.

[10] L. Chupin. Fokker-Planck in bounded domain. *Annales de l'Institut Fourier*, 60:217–255, 2010.

[11] Joseph A. Eichholz. *Discontinuous Galerkin methods for the radiative transfer equation and its approximaitions*. PhD thesis, The University of Iowa, 2011.

[12] L. C. Evans. *Partial Differential Equations*. American Mathematical Society, August 2008.

[13] S. J. Glick. Breast CT. *Annual Review of Biomedical Engineering*, 9:501–526, 2007.

[14] S. J. Glick, S. Thacker, X. Gong, and B. Liu. Evaluating the impact of X-rays spectral shape on image quality in at-panel CT breast imaging. *Medical Physics*, 34:5–24, 2007.

[15] T. Goldstein and S. Osher. The split bregman method for L1-regularized problems. *SIAM J. Imaging Sci.*, 2:323–343, 2009.

[16] P. González-Rodríguez and A. D. Kim. Light propagation in tissues with forward-peaked and large-angle scattering. *Applied Optics*, 47:2599–2609, 2008.

[17] J. Guinez and A. D. Rueda. Steady states for a Fokker-Planck equation on S_n. *Acta Math. Hungar*, 94:211–221, 2002.

[18] W. Han, J. Eichholz, X.-L. Cheng, and G. Wang. A theoretical framework of X-ray dark-field tomography. *SIAM J. Applied Math*, 71:1557–1577, 2011.

[19] W. Han, J. Eichholz, J. Huang, and J. Lu. RTE based bioluminescence tomography: a theoretical study. *Inverse Problems in Science and Engineering*, 19:435–459, 2011.

[20] W. Han, J. Eichholz, and G. Wang. On a family of differential approximations of the radiative transfer equation. *Journal of Mathematical Chemistry*, 50:689–7023, 2011.

[21] W. Han, J. Eichholz, and G. Wang. On a family of differential approximations of the radiative transfer equation. *Journal of Mathematical Chemistry*, 50:689–702, 2012.

[22] W. Han, J. A. Eichholz, and Q. Sheng. Theory of differential approximations of radiative transfer equation. In *Advances in Applied Mathematics and Approximation Theory 2012*, Springer Proceedings in Mathematics. Springer, New York, 2013.

[23] W. Han, J. Huang, and J. A. Eichholz. Discrete-ordinate discontinuous Galerkin methods for solving the radiative transfer equation. *SIAM J. Sci. Comput.*, 32:477–497, 2010.

[24] L. G. Henyey and J. L. Greenstein. Diffuse radiation in the galaxy. *Astrophys. J*, 20, January 1941.

[25] J. H. Joseph, W. J. Wiscombe, and J. A. Wienman. The delta-Eddington approximation for radiative flux transfer. *J. Atmos. Sci.*, 33:2452–2459, 1976.

[26] E. W. Larsen. The linear Boltzmann equation in optically thick systems with forward-peaked scattering. *Prog. Nucl. Energy*, 34:413–423, 1999.

[27] C. L. Leakeas and E. W. Larsen. Generalized Fokker-Planck approximations of particle transport with highly forward-peaked scattering. *Nucl. Sci. Eng.*, 137:236–250, 2001.

[28] E. E. Lewis and W. F. Miller. *Computational Methods of Neutron Transport.* John Wiley & Sons, New York, 1984.

[29] R. A. Lewis, K. D. Rogers, C. J. Hall, E. Towns-Andrews E, S. Slawson, A. Evans, S. E. Pinder, I. O. Ellis, C. R. Boggis, A. P. Hufton, and D. R. Dance. Breast cancer diagnosis using scattered X-rays. *Journal of Synchrotron Radiation*, 7:348–352, 2000.

[30] M. F. Modest. *Mathematical Methods in Image Reconstruction.* SIAM, Philadelphia, 2001.

[31] A. Momose, T. Takeda, Y. Itai, and K. Hirano. Phase-contrast X-ray computed tomography for observing biological soft tissues. *Nature Medicine*, 2:473–475, 1996.

[32] M. Muttarak, S. Pojchamarnwiputh, and B. Chaiwun. Efficacy of screening mammography–a meta analysis. *JAMA-Journal of the American Medical Association*, pages 149–154, 1995.

[33] M. Muttarak, S. Pojchamarnwiputh, and B. Chaiwun. Breast cancer in women under 40 years: Preoperative detection by mammography. *Annals Academy of Medicine Singapore*, 32:433–437, 2003.

[34] A. I. Noarov. Generalized solvability of the stationary Fokker-Planck equation. *Differ. Uravn.*, 43:813–819, 2007.

[35] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Model. Simul.*, 4:460–489, 2005.

[36] F. Pfeiffer, M. Bech, O. Bunk, P. Kraft, E. F. Eikenberry, Ch. Brönnimann, C. Grünzweig, and C. David. Hard X-ray dark-field imaging using a grating interferometer. *Nature Materials*, 7:134–137, 2008.

[37] F. Pfeiffer, O. Bunk, C. David, M. Bech, G. Le Duc, A. Bravin, and P. Cloetens. High-resolution brain tumor visualization using three-dimensional X-ray phase contrast tomography. *Physics in Medicine and Biology*, 52:6923–6930, 2007.

[38] E. D. Pisano, C. Gatsonis, E. Hendrick, M. Yaffe, J. K. Baum, S. Acharyya, E. F. Conant, L. L. Fajardo, L. Bassett, C. D'Orsi, R. Jong, and M. Rebner. Diagnostic performance of digital versus film mammography for breast-cancer screening. *Annals Academy of Medicine Singapore*, 353:433–437, 2005.

[39] G. C. Pomraning. The Fokker-Planck operator as an asymptotic limit. *Math. Models Methods Appl. Sci.*, 2:21–36, 1992.

[40] G. C. Pomraning. Higher order Fokker-Planck operators. *Nucl. Sci. Eng.*, 124:390–397, 1996.

[41] K. Przybylski and J. Ligou. Numerical analysis of the Boltzmann equation including Fokker-Planck terms. *Nucl. Sci. Eng.*, 81:92–109, 1982.

[42] I. Pucci-Minafra, C. Luparello, M. Andrioloand L. Basiricò, A. Aquino, and S. Minafra. A new form of tumour and fetal collagen that binds laminin. *Biochemistry*, 32:7421–7427, 1993.

[43] G. E. Thomas and K. Stamnes. *Radiative Transfer in the Atmosphere and Ocean.* Cambridge University Press, 1999.

[44] W. Zdunkowski, T. Trautmann, and A. Bott. *Radiation in the Atmosphere: A Course in Theoretical Meteorology.* Cambridge University Press, 2007.

[45] E. C. Zeeman. Stability of dynamical systems. *Nonlinearity*, 1:115–155, 1988.