Hindawi Publishing Corporation
Computational and Mathematical Methods in Medicine
Volume 2017, Article ID 2953560, 15 pages
http://dx.doi.org/10.1155/2017/2953560

Research Article

Mixed Total Variation and L^1 Regularization Method for Optical Tomography Based on Radiative Transfer Equation

Jinping Tang,[1] Bo Han,[1] Weimin Han,[2] Bo Bi,[3] and Li Li[1]

[1]*Department of Mathematics, Harbin Institute of Technology, Harbin, Heilongjiang Province, China*
[2]*Department of Mathematics, University of Iowa, Iowa City, IA, USA*
[3]*School of Mathematics and Statistics, Northeast Petroleum University, Daqing, Heilongjiang Province, China*

Correspondence should be addressed to Bo Han; bohan@hit.edu.cn

Received 22 March 2016; Accepted 27 September 2016; Published 9 February 2017

Academic Editor: Chuangyin Dang

Optical tomography is an emerging and important molecular imaging modality. The aim of optical tomography is to reconstruct optical properties of human tissues. In this paper, we focus on reconstructing the absorption coefficient based on the radiative transfer equation (RTE). It is an ill-posed parameter identification problem. Regularization methods have been broadly applied to reconstruct the optical coefficients, such as the total variation (TV) regularization and the L^1 regularization. In order to better reconstruct the piecewise constant and sparse coefficient distributions, TV and L^1 norms are combined as the regularization. The forward problem is discretized with the discontinuous Galerkin method on the spatial space and the finite element method on the angular space. The minimization problem is solved by a Jacobian-based Levenberg-Marquardt type method which is equipped with a split Bregman algorithms for the L^1 regularization. We use the adjoint method to compute the Jacobian matrix which dramatically improves the computation efficiency. By comparing with the other imaging reconstruction methods based on TV and L^1 regularizations, the simulation results show the validity and efficiency of the proposed method.

1. Introduction

Optical tomography with near-infrared light is a promising technique for noninvasive studying of the functional characters of human tissues. One can find many applications of this technique, for example, the early detection of breast cancer, cervical cancer screening, monitoring of infant brain tissue oxygenation level and functional brain activation, and the study of dosimetry in photon dynamic therapy; see [1–4]. Unlike the X-ray tomography, optical tomography uses the near-infrared (NIR) light as the probing radiation. Different human tissues have different absorption and scattering properties which affect the transmission of the NIR light. Therefore, we can identify the location and quantity of abnormal tissues by the emerging light measured on the boundary.

Optical tomography is a high contrast imaging modality. However, currently only low resolution reconstructions are possible, especially when using an unmodulated continuous wave source [5]. Another barrier of optical tomography is that reconstructed images have a poor quality particularly when abnormal targets are located deep in tissues. On the other hand, due to the limited measurement data, as well as the high scattering and high absorption properties, the identification problem of optical tomography usually is ill-posed and underdetermined. The ill-posedness makes the coefficient reconstruction sensitive to small perturbation from measurements, such as noise and computational error. Therefore, various reconstruction algorithms based on regularization are developed to obtain reasonable and stable reconstructions [6–9]. Recently, the regularization method with L^1 norm has been used in optical tomography [10, 11]. In [10], Levenberg-Marquardt strategy is applied for solving the L^2 regularization step of split Bregman algorithm. The contrast tests show the superiority of L^1 regularization over L^2 regularization. The numerical experiments therein also show that the L^1 regularization has a better utility when the independent measurements are much more limited. In [11], linearized Bregman iteration based on the Bregman distance is exploited to minimize the sparse regularization.

The experimental results in numerical simulation of an *in vivo* mouse demonstrate the effectiveness of the algorithm. However, the sparsity regularization may oversparsify the distribution of the coefficients.

Among many types of regularizations, TV regularization is often used due to its abilities of well reconstructing the discontinuous and piecewise constant distributions. TV regularization was first introduced in the field of image processing and image reconstruction; see [12–14] and the references therein. To date, TV regularization has been the preferred regularization strategy of considerable recent works, like photoacoustic tomography (PAT) [15–17], bioluminescence tomography (BLT) [18, 19], fluorescence tomography (FT) [20], optical coherence tomography (OCT) [21], and so on. In [18], the split Bregman algorithm for TV regularization (SBRTV) is applied to the reconstruction of source distribution in BLT. The algorithm is evaluated with 2D and 3D simulations and 3D *in vivo* experiments. In the reported results, the source distribution can be reconstructed with better accuracy for the location with SBRTV regularization than with L^2 or L^1 regularizations. In [19], the synergistic combination of Bregman method and TV regularization is utilized to quantitatively improve the reconstruction of absorption and scattering coefficients for both Jacobian-based and gradient-based methods in quantitative PAT. The feasibility of the algorithm is checked with 3D simulations.

For optical tomography, the coefficients are usually taken as piecewise constant. Since TV regularization is effective for piecewise constant reconstruction, it is a natural choice for reconstructing coefficients of piecewise constant distributions in optical tomography. However, due to the nonsmoothness and nondifferentiability, TV regularization is difficult to realize computationally. TV regularization for optical tomography based on the radiative transport equation has been studied in [22], where inexact Gauss-Newton method is used to solve the TV regularization problem. In [23], $W^{1,2}$ norm is chosen as the penalty in the regularization strategy for optical tomography based on the frequency radiative transport equation.

In this paper, we use regularization as a combination of TV and L^1 norm. To reduce the computational complexity, we apply the split Bregman method to solve the joint regularization problem. It is reasonable to combine L^1 regularization and TV regularization to improve the reconstruction quality in optical tomography. Moreover, based on the split Bregman method, the iterative algorithms can be computationally efficient. Various experiments in 2D are performed to evaluate the performance of the algorithm.

The rest of this paper is organized as follows. In Section 2, we briefly describe the forward and inverse problems for optical tomography. In Sections 3 and 4, we show the implementation details about computing the Fréchet derivative of the forward operator and the iterative procedure. In Section 5, numerical results are presented.

2. Optical Tomography

We consider two mathematical problems: the forward problem and the inverse problem. For the forward problem, based on the physical model of light propagation in tissues, for a given set of optical properties, we model the measurements on the boundary. For the inverse problem, the optical properties can be reconstructed by matching the predictions calculated from the forward problem and the measurements from the detectors.

2.1. Forward Problem. The light propagation in tissues is described by the radiative transport equation

$$\begin{aligned}&\boldsymbol{\omega}\cdot\nabla u(\mathbf{x},\boldsymbol{\omega})+\left(\mu_a(\mathbf{x})+\mu_s(\mathbf{x})\right)u(\mathbf{x},\boldsymbol{\omega})\\&\quad=\mu_s(\mathbf{x})\int_\Omega k(\boldsymbol{\omega}\cdot\widehat{\boldsymbol{\omega}})\,u(\mathbf{x},\widehat{\boldsymbol{\omega}})\,d\sigma(\widehat{\boldsymbol{\omega}})\quad\text{in } X\times\Omega.\end{aligned}\tag{1}$$

Here, $X\subset\mathbb{R}^d$, $d=2$ or 3, denotes a bounded convex domain with a C^1 boundary ∂X and $\Omega:=\mathbb{S}^{d-1}$ denotes the unit sphere of $\mathbb{R}^d$. The variables $\mathbf{x}$ and $\boldsymbol{\omega}$ denote the spatial position and the angular direction. $u(\mathbf{x},\boldsymbol{\omega})$ describes the density of photons. The expression $\boldsymbol{\omega}\cdot\nabla u(\mathbf{x},\boldsymbol{\omega})$ denotes the directional derivative at position $\mathbf{x}$ along the direction $\boldsymbol{\omega}$. The nonnegative normalized phase function $k(\boldsymbol{\omega}\cdot\widehat{\boldsymbol{\omega}})$ is the probability that photons traveling in the direction $\widehat{\boldsymbol{\omega}}$ are scattered into the direction $\boldsymbol{\omega}$. In optical tomography, the phase function usually is taken as the Henyey-Greenstein phase function (cf. [24]): in two dimensions, it is of the form

$$k(\boldsymbol{\omega}\cdot\widehat{\boldsymbol{\omega}})=\frac{1-g^2}{2\pi\left(1+g^2-2g\boldsymbol{\omega}\cdot\widehat{\boldsymbol{\omega}}\right)},\tag{2}$$

where the parameter $g\in(-1,1)$ is the anisotropy factor of the scattering medium. The absorption coefficient is denoted by $\mu_a(\mathbf{x})$, and the scattering coefficient is denoted by $\mu_s(\mathbf{x})$. For the optical parameters $\mu_a(\mathbf{x})$ and $\mu_s(\mathbf{x})$, the following conditions are assumed to hold throughout this presentation.

Assumptions

(A1) The function $\mu_a(\mathbf{x})$ is uniformly positive and bounded; that is, there exist two positive constants μ_a^1 and μ_a^2 such that $0<\mu_a^1\le\mu_a\le\mu_a^2<\infty$ a.e. in X.

(A2) The function $\mu_s(\mathbf{x})$ is uniformly positive and bounded; that is, there exist two positive constants μ_s^1 and μ_s^2 such that $0<\mu_s^1\le\mu_s\le\mu_s^2<\infty$ a.e. in X.

Equation (1) is supplemented by boundary conditions. Similar to the X-ray CT, optical tomography experiments acquire the current distribution of detectors on the boundary under multi-incidents. Let ζ_i, $1\le i\le s$, be disjoint, connected subsets of ∂X. Corresponding to s incident sources $u_{\text{in},i}$ on ζ_i, $1\le i\le s$, define $u_i(\mathbf{x},\boldsymbol{\omega})$ by

$$\begin{aligned}&\boldsymbol{\omega}\cdot\nabla u_i(\mathbf{x},\boldsymbol{\omega})+\left(\mu_a(\mathbf{x})+\mu_s(\mathbf{x})\right)u_i(\mathbf{x},\boldsymbol{\omega})\\&\quad=\mu_s(\mathbf{x})\int_\Omega k(\boldsymbol{\omega}\cdot\widehat{\boldsymbol{\omega}})\,u_i(\mathbf{x},\widehat{\boldsymbol{\omega}})\,d\sigma(\widehat{\boldsymbol{\omega}})\quad\text{in } X\times\Omega,\end{aligned}$$

$$u_i(\mathbf{x},\boldsymbol{\omega}) = \begin{cases} u_{\text{in},i}(\mathbf{x},\boldsymbol{\omega}), & \mathbf{x}\in\zeta_i,\ \boldsymbol{\omega}\cdot\boldsymbol{\nu}(\mathbf{x})<0, \\ 0, & \text{otherwise}, \end{cases} \quad \text{on } \partial X\times\Omega, \tag{3}$$

where $\boldsymbol{\nu}(\mathbf{x})$ is the unit outward normal vector at $\mathbf{x}\in\partial X$.

Corresponding to each independent incident $u_{\text{in},i}(\mathbf{x},\boldsymbol{\omega})$, the measurable quantity is the outgoing light $M_i(\mathbf{x})$ on the boundary of domain, which can be written as

$$M_i(\mathbf{x}) = \int_{\boldsymbol{\omega}\cdot\boldsymbol{\nu}(\mathbf{x})>0} \boldsymbol{\omega}\cdot\boldsymbol{\nu}(\mathbf{x})\, u_i(\mathbf{x},\boldsymbol{\omega})\, d\sigma(\boldsymbol{\omega}), \quad \mathbf{x}\in\partial X. \tag{4}$$

If we assume there are $\mathfrak{d}$ detectors located at $\mathfrak{d}$ different positions $\xi_j, 1 \le j \le \mathfrak{d}$, then the optical tomography experiment consists of exciting the domain X with a sequence of incident source $u_{\text{in},i}$ and recording the corresponding measurements data $M_{j,i} = M_i(\mathbf{x})|_{\mathbf{x}=\xi_j}$. Here, j and i represent the row index and column index in matrix $M_{j,i}$, respectively.

With above notations, a mathematical description of such an experiment is provided by the following nonlinear forward operator:

$$\begin{aligned} F_i &: D \longrightarrow L^2(\partial X), \\ (\mu_a,\mu_s) &\longmapsto M_i(\mathbf{x}), \\ & 1\le i\le s, \end{aligned} \tag{5}$$

which maps prescribed optical parameters to the corresponding measurements data. Here, F_i denotes the ith forward operator corresponding to the ith incident source and the resulting measurement data on $\mathfrak{d}$ detectors. The forward operator F_i is well defined for μ_a and μ_s in the set

$$D = \{(\mu_a,\mu_s) \text{ satisfying assumptions (A1)-(A2)}\} \tag{6}$$

(cf. [25]).

There are many references on the discretization of RTE; see, for instance, [26–30]. In this paper, we use continuous linear elements and discontinuous Galerkin method with piecewise linear functions to discretize the angular variables and the spatial variables, respectively. An upwind numerical flux is used to approximate the incoming flux through the surface of the control element and inflow boundary. After assembling the full discretization formulation and forming a system of tebra equation, we solve the linear system by Gauss-Seidel method. We use piecewise constants to approximate the absorption and scattering coefficients. We can express (μ_a,μ_s) as

$$\begin{aligned} \mu_a(\mathbf{x}) &\approx \sum_{k=1}^{N} \mu_{a,k}\chi_k(\mathbf{x}), \\ \mu_s(\mathbf{x}) &\approx \sum_{k=1}^{N} \mu_{s,k}\chi_k(\mathbf{x}), \end{aligned} \tag{7}$$

where N denotes the number of the elements, $\chi_k(\mathbf{x})$ denotes the character function corresponding to the kth element, and $\mu_{a,k}$ and $\mu_{s,k}$ denote the values of absorption and scattering coefficients on the kth element.

2.2. Inverse Problem. The inverse problem of optical tomography is to determine the unknown coefficients μ_a and μ_s from the boundary detector readings. In this paper, we only reconstruct the absorption coefficient assuming that the scattering coefficient is known; then the forward operator F_i in fact acts on the unknown μ_a only. Thus, the inverse problem of optical tomography is to determine μ_a from the following system of nonlinear equations:

$$F_i(\mu_a) = M_i, \quad 1\le i\le s. \tag{8}$$

Then the optical tomography can be formulated as minimizing the difference between the measurement data and the model predictions

$$\mu_a = \arg\min_D \frac{1}{2}\sum_{i=1}^{s} \left\| F_i(\mu_a) - M_i \right\|^2_{L^2(\partial X)}, \tag{9}$$

known as data fidelity. The inverse problem (9) is ill-posed in the sense that the amount of the measurements is quite limited compared with the number of the unknowns and that the measurements contain noises. To overcome the ill-posedness, the data fidelity should be combined with appropriate regularization. In the regularization strategy, we minimize the following objective functional:

$$\mu_a = \arg\min_D \left\{ \frac{1}{2}\sum_{i=1}^{s} \left\| F_i(\mu_a) - M_i \right\|^2_{L^2(\partial X)} + \alpha R(\mu_a) \right\}. \tag{10}$$

Here, $R(\mu_a)$ is the regularization penalty functional that enforces *a prior* information on μ_a and $\alpha > 0$ is the regularization parameter that trades off the weight between the discrepancy term and the penalty functional.

According to different ways of treating the derivative of data fidelity, there are two sorts of approaches for solving (10). One is to linearize the forward operator F_i near the nth iteration of μ_a, which is denoted as μ_a^n, as follows:

$$F_i(\mu_a) \approx F_i(\mu_a^n) + F_i'(\mu_a^n)(\mu_a - \mu_a^n), \tag{11}$$

where $F_i'(\mu_a^n)$ is the derivative of the forward operator F_i with respect to μ_a^n. The linearized formulation provides a good approximation when μ_a^n is close to the true value. As a result, the minimized problem (10) is treated by an iterative procedure as follows:

$$\begin{aligned} \mu_a^{n+1} = \arg\min_D \Bigg\{ & \frac{1}{2} \\ & \cdot \sum_{i=1}^{s} \left\| F_i'(\mu_a^n)(\mu_a - \mu_a^n) + F_i(\mu_a^n) - M_i \right\|^2_{L^2(\partial X)} \\ & + \alpha R(\mu_a^n) \Bigg\}. \end{aligned} \tag{12}$$

This Jacobian-based minimization problem can be solved with many iteratively optimization techniques, such as the Levenberg-Marquardt type method (LM). LM type method is

the special case of Gauss-Newton type method. The standard LM method (or Gauss-Newton method) uses $\|\mu_a - \mu_a^n\|^2_{L^2(X)}$ as the penalty. The LM type method defines the update μ_a^{n+1} in a region around μ_a^n, while the Gauss-Newton method always defines μ_a^{n+1} in a neighbourhood of the initial guess μ_a^0 for each $n \geq 0$. Hence, from the optimization point of view, LM type method is more favourable in nature. Based on the Jacobian-based method, many optimization techniques such as split Bregman method can be used to solve (12). On the other hand, when μ_a^n is close to the true value, the Jacobian-based method shares the typical quadratic convergence from Newton method.

Alternatively, (10) can be solved directly by minimizing the nonlinear functional with some gradient-based methods, like Quasi-Newton method, nonlinear conjugate gradient method, limited-memory BFGS method, and so on. Thus, the computation of the linearized Jacobian matrix is avoided and only the gradient of the nonlinear minimizing functional needs to be computed. The gradient-based method in general has superlinear convergence and is economic in memory storage which is suitable for large scale problems, such as 3D problems. Since our numerical experiments are all done in two dimensions and the problem scale is not so large as that in three dimensions, we use the Jacobian-based LM type method to solve (10).

3. Adjoint Problem

Since we adopt the Jacobian-based minimization method to solve (10) in this paper, the derivative of the forward operator becomes a matrix which is called Jacobi matrix. For convenience of representing the element of Jacobi matrix hereinafter, we use $[J^i_{\mu_a}]$ to represent the Jacobi matrix of F_i, with respect to $\mu_a = (\mu_{a,1}, \ldots, \mu_{a,N})$, $1 \leq i \leq s$; here $\mu_{a,i}$ denotes the ith element of the discretized μ_a. For each i, $[J^i_{\mu_a}] \in \mathbb{R}^{\mathfrak{d} \times N}$ is defined as follows:

$$\left[J^i_{\mu_a}\right] = \begin{pmatrix} \dfrac{\partial M_{1,i}}{\partial \mu_{a,1}} & \dfrac{\partial M_{1,i}}{\partial \mu_{a,2}} & \cdots & \dfrac{\partial M_{1,i}}{\partial \mu_{a,N}} \\ \dfrac{\partial M_{2,i}}{\partial \mu_{a,1}} & \dfrac{\partial M_{2,i}}{\partial \mu_{a,2}} & \cdots & \dfrac{\partial M_{2,i}}{\partial \mu_{a,N}} \\ \cdots & \cdots & \vdots & \cdots \\ \dfrac{\partial M_{\mathfrak{d},i}}{\partial \mu_{a,1}} & \dfrac{\partial M_{\mathfrak{d},i}}{\partial \mu_{a,2}} & \cdots & \dfrac{\partial M_{\mathfrak{d},i}}{\partial \mu_{a,N}} \end{pmatrix}, \quad 1 \leq i \leq s. \tag{13}$$

$[J^i_{\mu_a}]$ can be computed by direct method which differentiates $M_{\zeta_i}(\xi_j)$ with respect to the perturbation of μ_a on each element. In order to compute $[J^i_{\mu_a}]$, for each i, we need to solve forward problems $\mathfrak{d} \times N$ times. For a total of s sources, we need to compute $s \times \mathfrak{d} \times N$ times, so the direct method is very time consuming. Therefore, we use the adjoint formulation to compute $[J^i_{\mu_a}]$ instead of direct method.

We first consider the analytic Fréchet derivative of the forward operator for the boundary value problem of RTE given by [25].

$$F_i'(\mu_a)^* M_{\zeta_i}(\mathbf{x}) = -\int_\Omega u_i(\mathbf{x}, \boldsymbol{\omega})\, \varphi(\mathbf{x}, \boldsymbol{\omega})\, d\sigma(\boldsymbol{\omega}), \tag{14}$$

where $F_i'(\mu_a)^*$ denotes the adjoint operator of $F_i(\mu_a)$ and $\varphi(\mathbf{x}, \boldsymbol{\omega})$ is the solution of the adjoint RTE

$$\begin{aligned} &-\boldsymbol{\omega} \cdot \nabla \varphi(\mathbf{x}, \boldsymbol{\omega}) + (\mu_a + \mu_s)\, \varphi(\mathbf{x}, \boldsymbol{\omega}) \\ &\quad = \mu_s \int_\Omega k(\boldsymbol{\omega}, \widehat{\boldsymbol{\omega}})\, \varphi(\mathbf{x}, \widehat{\boldsymbol{\omega}})\, d\sigma(\widehat{\boldsymbol{\omega}}) \quad \text{in } X \times \Omega \end{aligned} \tag{15}$$

with boundary condition

$$\varphi(\mathbf{x}, \boldsymbol{\omega}) = \begin{cases} (\boldsymbol{\omega} \cdot \boldsymbol{\nu}(\mathbf{x}))\, M_i(\mathbf{x}), & \mathbf{x} \in \partial X,\ \boldsymbol{\omega} \cdot \boldsymbol{\nu}(\mathbf{x}) > 0, \\ 0, & \text{otherwise}, \end{cases} \quad \text{on } \partial X \times \Omega. \tag{16}$$

From the adjoint RTE (15) and the boundary condition (16), it seems that the adjoint problem should be solved in the reverse direction of propagation with a completely new computation solver. However, by the standard reciprocity theorem for the Boltzmann equation given in [31],

$$G(\mathbf{x}, \boldsymbol{\omega}; \mathbf{x}_0, \boldsymbol{\omega}_0) = G(\mathbf{x}_0, -\boldsymbol{\omega}_0; \mathbf{x}, -\boldsymbol{\omega}), \tag{17}$$

which states that the angular density at $\mathbf{x}$ in direction $\boldsymbol{\omega}$ due to a source at $\mathbf{x}_0$ in direction $\boldsymbol{\omega}_0$ is the same as the angular density at $\mathbf{x}_0$ in direction $-\boldsymbol{\omega}_0$ due to a source at $\mathbf{x}$ in direction $-\boldsymbol{\omega}$. Here, $G(\cdot;\cdot)$ presents the Green function of (1) for isotropic point source on the boundary. Then the adjoint problems (15) and (16) can be transformed into the same form as the forward problem, simply replacing direction $\boldsymbol{\omega}$ with $-\boldsymbol{\omega}$. Then we just need to solve the following equation for the radiance $\widehat{\varphi}(\mathbf{x}, -\boldsymbol{\omega}) = \varphi(\mathbf{x}, \boldsymbol{\omega})$ by the same forward solver:

$$\begin{aligned} &\boldsymbol{\omega} \cdot \nabla \widehat{\varphi}(\mathbf{x}, \boldsymbol{\omega}) + (\mu_a + \mu_s)\, \widehat{\varphi}(\mathbf{x}, \boldsymbol{\omega}) \\ &\quad = \mu_s \int_\Omega k(\boldsymbol{\omega}, \widehat{\boldsymbol{\omega}})\, \widehat{\varphi}(\mathbf{x}, \widehat{\boldsymbol{\omega}})\, d\sigma(\widehat{\boldsymbol{\omega}}), \end{aligned} \tag{18}$$

with adjoint boundary condition

$$\widehat{\varphi}(\mathbf{x}, \boldsymbol{\omega}) = \begin{cases} (\boldsymbol{\omega} \cdot \boldsymbol{\nu}(\mathbf{x}))\, M_i(\mathbf{x}), & \mathbf{x} \in \partial X,\ \boldsymbol{\omega} \cdot \boldsymbol{\nu}(\mathbf{x}) < 0, \\ 0, & \text{otherwise}. \end{cases} \tag{19}$$

This means we need to solve (18) and (19) with the forward solver firstly and then reverse all directions on solution.

If we consider one source position ζ_i and one detector position ξ_j, then the kth column of the Jacobi matrix $[J^i_{\mu_a}]$ can be computed as follow:

$$\left[J^i_{\mu_a}\right](:,k) = \int_X \chi_k(\mathbf{x}) \int_\Omega u_i(\mathbf{x}, \boldsymbol{\omega})\, \varphi_j(\mathbf{x}, \boldsymbol{\omega})\, d\sigma(\boldsymbol{\omega})\, d\mathbf{x}, \tag{20}$$

where φ_j is the solution of the following boundary value problem:

$$\begin{aligned} &-\boldsymbol{\omega}\cdot\nabla\varphi_j(\mathbf{x},\boldsymbol{\omega})+\left(\mu_a(\mathbf{x})+\mu_s(\mathbf{x})\right)\varphi_j(\mathbf{x},\boldsymbol{\omega})\\ &\quad=\mu_s(\mathbf{x})\int_\Omega k(\boldsymbol{\omega}\cdot\widehat{\boldsymbol{\omega}})\,\varphi_j(\mathbf{x},\widehat{\boldsymbol{\omega}})\,d\sigma(\widehat{\boldsymbol{\omega}}),\\ &\varphi_j(\mathbf{x},\boldsymbol{\omega})=\begin{cases}(\boldsymbol{\omega}\cdot\boldsymbol{\nu}(\mathbf{x}))\,M_i(\mathbf{x})\big|_{\mathbf{x}=\xi_j}, & \boldsymbol{\omega}\cdot\boldsymbol{\nu}(\mathbf{x})>0,\\ 0, & \text{otherwise.}\end{cases}\end{aligned} \tag{21}$$

Thus for the adjoint method to compute $[J^i_{\mu_a}]$, for each $1\le i\le s$, we only need to solve forward problems $\mathfrak{d}$ times. For s sources, we need to solve forward problems $s\times\mathfrak{d}$ times, which dramatically reduces the computation burden and improves the efficiency of the algorithm.

4. Iterative Procedure

The unknown μ_a can be estimated through the regularization method. The quality of reconstructed image strongly depends on the choice of the penalty term $R(\mu_a)$. If we choose the L^2 norm as penalty, that is, $R(\mu_a)=\|\mu_a\|_2^2$, the reconstructed image usually blurs with a low resolution. If we choose L^1 norm as the penalty, that is, $R(\mu_a)=\|\mu_a\|_1$, the reconstructed image tends to find a sparse solution [10, 18]. To treat the discontinuity and the edges of different distribution regions, total variation is usually chosen as the penalty functional, that is,

$$R(\mu_a)=\int_X|\nabla\mu_a|. \tag{22}$$

The symbol $\int_X|\nabla\mu_a|$ denotes the total variation seminorm [32] of $\mu_a\in L^1(X)$ as follows:

$$\begin{aligned}\int_X|\nabla\mu_a|:=\sup\Big\{&\int_X\mu_a\operatorname{div}\varphi\,dx\mid\varphi\\ &\in C_0^\infty\left(X;\mathbb{R}^d\right),\|\varphi\|_{L^\infty(X;\mathbb{R}^d)}\le1\Big\}.\end{aligned} \tag{23}$$

To improve the reconstruction quality, we consider the total variation mixed with the L^1 norm as the penalty. The functional to be minimized is of the form

$$\begin{aligned}J(\mu_a)=&\frac{1}{2}\sum_{i=1}^s\|F_i(\mu_a)-M_i\|^2_{L^2(\partial X)}\\ &+\alpha\int_X|\nabla\mu_a|+\beta\|\mu_a\|_{l_1},\end{aligned} \tag{24}$$

where α, β are the regularization parameters. We will apply the split Bregman method to solve (24) [33]. Instead of (24), we consider the following constrained optimization problem:

$$\begin{aligned}\inf_{(\mu_a,\mathscr{D})}&\frac{1}{2}\sum_{i=1}^s\|F_i(\mu_a)-M_i\|^2_{L^2(\partial X)}+\alpha\int_X|\nabla\mu_a|+\beta\|\mathscr{D}\|_{l_1}\\ &\text{such that }\mathscr{D}=\mu_a.\end{aligned} \tag{25}$$

Solution of the above minimization problem can be obtained by solving the unconstrained optimization problem

$$\begin{aligned}(\mu_a,\mathscr{D})=\arg\min&\frac{1}{2}\sum_{i=1}^s\|F_i(\mu_a)-M_i\|^2_{L^2(\partial X)}\\ &+\alpha\int_X|\nabla\mu_a|+\beta\|\mathscr{D}\|_{l_1}+\frac{\eta}{2}\|\mathscr{D}-\mu_a\|^2_{l^2},\end{aligned} \tag{26}$$

where $\eta>0$ is the split parameter. Now let us iteratively solve the following subproblems [34]:

$$\begin{aligned}&\left(\mu_a^{n+1},\mathscr{D}^{n+1}\right)\\ &\quad=\arg\min\frac{1}{2}\sum_{i=1}^s\|F_i(\mu_a)-M_i\|^2_{L^2(\partial X)}\\ &\qquad+\alpha\int_X|\nabla\mu_a|+\beta\|\mathscr{D}\|_{l_1}+\frac{\eta}{2}\|\mathscr{D}-\mu_a-b_d^n\|_2^2,\end{aligned} \tag{27}$$

with the following update for b_d:

$$b_d^{n+1}=b_d^n+\mu_a^{n+1}-\mathscr{D}^{n+1}. \tag{28}$$

The minimization of subproblems in (1) can be iteratively solved by splitting it into the minimizations of μ_a and $\mathscr{D}$ separately. This suggests the following steps.

Step 1.

$$\begin{aligned}\mu_a^{n+1}=\arg\min_{\mu_a}&\frac{1}{2}\sum_{i=1}^s\|F_i(\mu_a)-M_i\|^2_{L^2(\partial X)}+\alpha\int_X|\nabla\mu_a|\\ &+\frac{\eta}{2}\|\mathscr{D}^n-\mu_a-b_d^n\|_2^2.\end{aligned} \tag{29}$$

Step 2.

$$\mathscr{D}^{n+1}=\arg\min_{\mathscr{D}}\beta\|\mathscr{D}\|_{l_1}+\frac{\eta}{2}\left\|\mathscr{D}-\mu_a^{n+1}-b_d^n\right\|_2^2. \tag{30}$$

Step 3.

$$b_d^{n+1}=b_d^n+\mu_a^{n+1}-\mathscr{D}^{n+1}. \tag{31}$$

For the solution of Step 1, we use the Levenberg-Marquardt method, which has a high convergence rate. Thus we solve a minimization problem as follows.

*Step 1**

$$\begin{aligned}\mu_a^{n+1}=\arg\min_{\mu_a}&\frac{1}{2}\sum_{i=1}^s\left\|F_i(\mu_a^n)+\left[J^i_{\mu_a^n}\right](\mu_a-\mu_a^n)-M_i\right\|_2^2\\ &+\alpha\int_X|\nabla\mu_a|+\frac{\eta}{2}\|\mathscr{D}^n-\mu_a-b_d^n\|_2^2,\end{aligned} \tag{32}$$

where $[J^i_{\mu_a^n}]$ denotes the Jacobi matrix of the forward operator F_i with respect to μ_a^n. For solving Step 1*, in order to avoid the nondifferentiability of the total variation term $\int_X|\nabla\mu_a|$

at a zero point, we approximate it with a smooth functional $\|\mu_a\|_{\mathrm{TV},\varepsilon}$ defined as

$$\|\mu_a\|_{\mathrm{TV},\varepsilon} = \int_X \sqrt{|\nabla\mu_a|^2 + \epsilon^2}\, dx, \quad \epsilon > 0. \tag{33}$$

The parameter ϵ is a constant, and it cannot be too large or too small. The Euler-Lagrange equation for Step 1^* is

$$\begin{aligned}\sum_{i=1}^{s} \left[J^i_{\mu_a^n}\right]^T \left(\left[J^i_{\mu_a^n}\right](\mu_a - \mu_a^n) - R_i(\mu_a^n)\right) + \alpha L(\mu_a^n)\mu_a \\ + \eta(\mathscr{D}^n - \mu_a - b_d^n) = 0,\end{aligned} \tag{34}$$

where $[\cdot]^T$ denotes the transpose, $R_i(\mu_a^n)$ is the short notion of $F_i(\mu_a^n) - M_i$, the term $L(\mu_a^n)\mu_a$ is defined as

$$L(\mu_a^n)\mu_a = -\nabla\cdot\left\{\frac{\nabla\mu_a}{\sqrt{|\nabla\mu_a^n|^2 + \epsilon^2}}\right\}. \tag{35}$$

Then we can get the iterative update by solving the above equation with Newton method as follows:

$$\begin{aligned}&\left(\sum_{i=1}^{s}\left[J^i_{\mu_a^n}\right]^T\left[J^i_{\mu_a^n}\right] + \alpha L(\mu_a^n) + \eta I\right)(\mu_a - \mu_a^n) \\ &= \sum_{i=1}^{s}\left[J^i_{\mu_a^n}\right]^T R_i(\mu_a^n) + \alpha L(\mu_a^n)\mu_a^n \\ &\quad + \eta(\mathscr{D}^n - \mu_a^n - b_d^n).\end{aligned} \tag{36}$$

Step 2 is an L^1 norm regularization problem and it can be solved efficiently through the shrinkage operator, that is,

$$\mathscr{D}^{n+1} = \mathrm{shrink}\left(\mu_a^{n+1} + b_d^n, \frac{\beta}{\eta}\right), \tag{37}$$

where the shrinkage operator

$$\begin{aligned}\mathrm{shrink}(x,t) &= \sin(x)\max(|x| - t, 0) \\ &= \begin{cases} x - t, & x \geq t, \\ 0, & |x| < t, \\ x + t, & x \leq -t. \end{cases}\end{aligned} \tag{38}$$

Implementation of the split Bregman formulation for our TV-L^1 regularization is described in Algorithm 1.

5. Reconstruction Results and Discussion

In this section, the TV-L^1 regularization with the split Bregman formulation is applied to 2D test problems. We assume the distribution of the scattering coefficient is known. Reconstructions of spatially dependent distributions of absorption coefficient inside the medium are performed and discussed. In our simulation, a circular domain X which contains different inclusions is investigated. The radius of the circle is 10 mm. 12 sources and 12 detectors are located on the boundary of the domain with equal space. This yields totally 144 source-detector pairs to be used in the inversion.

Noise-free synthetic data are generated by solving the forward problem on triangular meshes with the method we mentioned in Section 2.1. Note that the meshes for the inverse problem are coarser than the meshes for the forward problem in order to avoid the "inverse crime" [35] and for the purpose of testing the robustness of the proposed algorithm. We will describe the meshes in each case. In all the cases, the angular space is discretized into 32 directions which equally divide the interval $[0, 2\pi]$.

The background medium has scattering coefficient of $10\,\mathrm{mm}^{-1}$ and absorption coefficient of $0.01\,\mathrm{mm}^{-1}$, respectively, and these values keep the same throughout this paper. The initial guess of reconstruction is set to be identical to the properties of the background medium. That is, the iterative procedure started with the background value of the absorption coefficient.

The incident impulse on the inflow boundary is settled as

$$\begin{aligned}u_{\mathrm{in},i}(\mathbf{x},\boldsymbol{\omega}) &= B_{i,\omega_i}, \\ (\mathbf{x},\boldsymbol{\omega}) &\in \partial X \times \Omega, \\ \boldsymbol{\omega}\cdot\nu(\mathbf{x}) &< 0, \\ 1 \leq i &\leq s,\end{aligned} \tag{39}$$

where B_{i,ω_i} is a piecewise linear function whose spatial support is S_i and achieves the value 1 at the center node of S_i, where S_i is the element through which the ith incident impulse passes. The direction of $u_{\mathrm{in},i}$ points approximately from the center of S_i to the center of X.

Based on above settings, we use various simulations to validate the proposed split Bregman algorithm for our TV-L^1 regularization. Our purpose is to show the following results.

First, by comparing the reconstruction with different anisotropic factors g, the proposed algorithm works better when g is bigger.

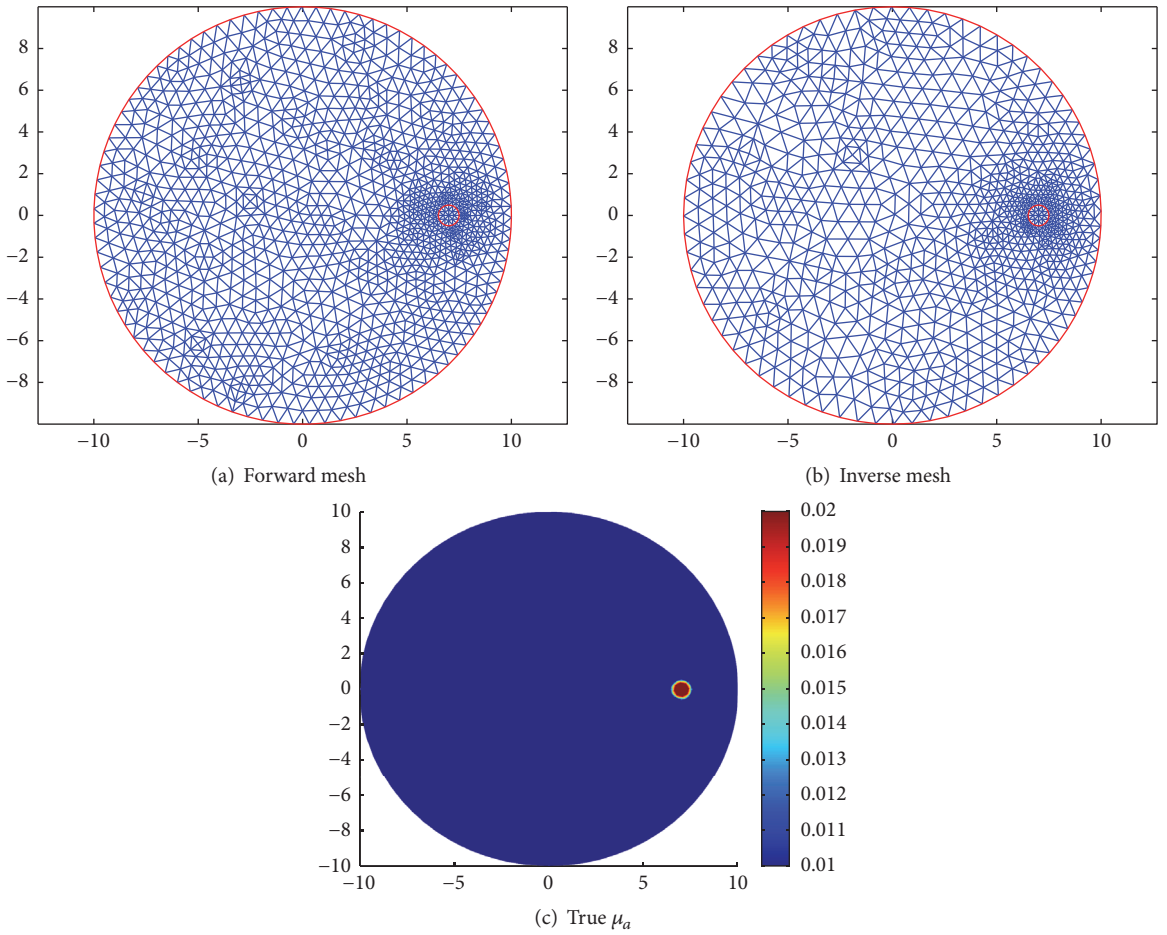
Second, for small sparse inclusions, the TV-L^1 regularization can reconstruct the location and the quality of the coefficient more accurately than the TV regularization and the L^1 regularization.

Third, the proposed split algorithm for TV-L^1 needs less computation time than the Levenberg-Marquardt algorithm for TV regularization. Even though our proposed algorithm needs more computation time than the split Bregman algorithm for L^1 regularization [10], the reconstruction quality with the proposed algorithm is much better than that with the split Bregman algorithm for L^1 regularization.

5.1. Simulation 1: Reconstruction with Different Anisotropic Factors. In this simulation, we reconstruct one circle inclusion with the radius 0.5 mm centered at (7.0 mm, 0.0 mm). In this circle, the absorption coefficient $\mu_a = 0.02\,\mathrm{mm}^{-1}$ and $\mu_s = 20\,\mathrm{mm}^{-1}$. Our goal is to reconstruct μ_a. For solving the inverse problem, we use a mesh of 772 nodes and 1484 triangular elements. To avoid the inverse crime, we use a

Input: Given the initial guess μ_a^0; the regularization parameter $\alpha > 0$, $\beta > 0$, $\eta > 0$;
$\mathscr{D}^0 = b_d^0 = 0$, and the tolerance ε.
while $\|\mu_a^{n+1} - \mu_a^n\| > \varepsilon$ **do**
For $1 \le i \le s$, compute μ_a^{n+1} with (36);
Compute $\mathscr{D}^{n+1} = \text{shrink}(\mu_a^{n+1} + b_d^n, \beta/\eta)$;
Compute $b_d^{n+1} = b_d^n + \mu_a^{n+1} - \mathscr{D}^{n+1}$.
end while
Output: Output an approximation $\mu_a = \mu_a^{n+1}$.

Algorithm 1: Reconstruction algorithm based on the split Bregman method for TV-L^1 regularization.



(a) Forward mesh (b) Inverse mesh

(c) True μ_a

Figure 1: The forward mesh, inverse mesh, and the true distribution of μ_a in simulation 1.

mesh of 1296 nodes and 2440 triangular elements for the synthetic data. The simulated absorption distributions and the meshes can be seen in Figure 1. We solve the minimizer of $J(\mu_a)$ by using the split Bregman method for TV-L^1 method with exact data, that is, $\delta = 0$, in this example. Then we compare the reconstruction results for different anisotropic factors g. For the comparison purpose, we use the same parameters α, β, η, and ε for different g. Noticing that the so-called exact data in fact contains noise, so the regularization parameter α, β should not be too small. Since the inclusion is very small in this example, we enhance more weight of L^1 penalty than the weight of TV penalty. Here we take $\alpha = 10^{-4}$ and $\beta = 10^{-3}$. The other two parameters are taken as $\eta = 10^{-5}$ and $\epsilon = 10^{-6}$, respectively.

The reconstruction results can be seen in Figure 2. The first row of Figure 2 is the reconstruction results for $g = 0.1$ and $g = 0.4$ from the left to the right. The second row of Figure 2 is the reconstruction results for $g = 0.7$ and $g = 0.9$ from the left to the right. As it can be seen, the reconstructions for the bigger g are more clear than that for

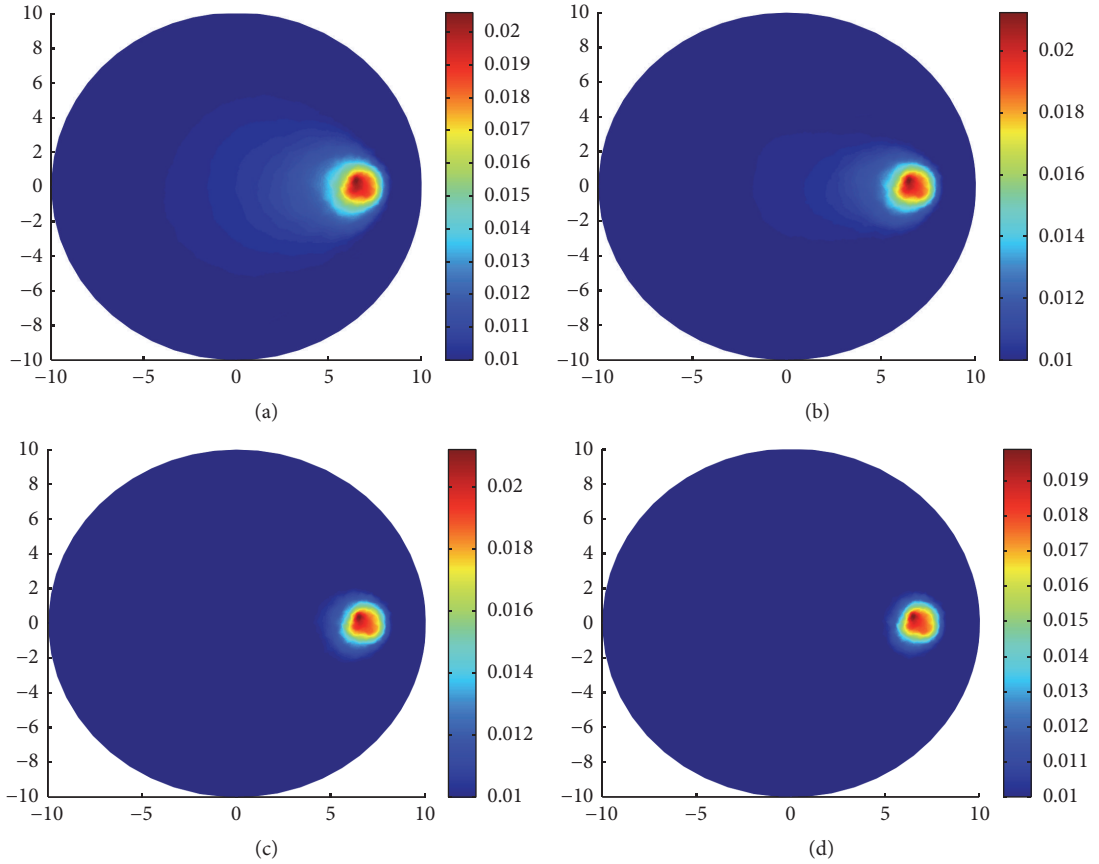


FIGURE 2: Reconstruction results with and without scaling strategy. (a) Reconstruction of μ_a for $g = 0.1$; (b) reconstruction of μ_a for $g = 0.4$; (c) reconstruction of μ_a for $g = 0.7$; (d) reconstruction of μ_a for $g = 0.9$. The reconstructions are done by using the split Bregman algorithm for TV-L^1 regularization with noise-free synthetic data.

smaller g. There are some apparent blurred parts in the left side of the inclusion in Figures 2(a) and 2(b). The blurred parts are much smaller in Figure 2(c) than in Figure 2(b) and almost disappeared in Figure 2(d).

From this example, we can find that the proposed algorithm works better for the bigger anisotropic factor g. The following simulations in this paper are all performed with $g = 0.9$.

5.2. Simulation 2: Reconstruction with TV-L^1 Regularization and the Standard TV Regularization. There are two groups of experiments in this section. In the first group of experiment, we design three types of inclusions to compare the performance of the TV-L^1 regularization and the TV regularization with noise-free synthetic data. In the first case, a small inclusion with radius 0.5 mm centered at (7.0 mm, 0.0 mm) is designed for $\mu_a = 0.02\,\text{mm}^{-1}$ and $\mu_s = 20.0\,\text{mm}^{-1}$, respectively. In the second case, a middle inclusion with radius 2 mm centered at (5.0 mm, 0.0 mm) is designed for $\mu_a = 0.02\,\text{mm}^{-1}$ and $\mu_s = 20\,\text{mm}^{-1}$, respectively. In the third case, a middle inclusion with radius 4 mm centered at (3.0 mm, 0.0 mm) is designed for $\mu_a = 0.02\,\text{mm}^{-1}$ and $\mu_s = 20\,\text{mm}^{-1}$, respectively. The simulated true absorption coefficient can be seen on the first column of Figure 3. For the small inclusion case, we use a mesh of 772 nodes and 1484 triangular elements for the inverse problem and a mesh of 1267 nodes and 2440 triangular elements for the forward problem, the simulated true absorption coefficient and reconstruction results with TV-L^1 regularization and TV regularization are shown in Figure 3. For the middle inclusion case, we use a mesh of 575 nodes and 1080 triangular elements for the inverse problem and a mesh of 813 nodes and 2642 triangular elements for the forward problem, the simulated true absorption coefficient and reconstruction results with TV-L^1 regularization and TV regularization are shown in Figure 4. For the big inclusion case, we use a mesh of 583 nodes and 1096 triangular elements for the inverse problem and a mesh of 871 nodes and 3124 triangular elements for the forward problem, the simulated true absorption coefficient and reconstruction results with TV-L^1 regularization and TV regularization are shown in Figure 5. The values of the parameters are shown in Table 1. Since we use the same $\varepsilon = 10^{-6}$ as the last section, we will not present it repeatedly in Table 1.

Observing from the results, we can see that, for small and middle inclusions, the TV-L^1 regularization performs better than the TV regularization in both localizing the location and quantifying the values; see Figures 3(b) and 4(b)

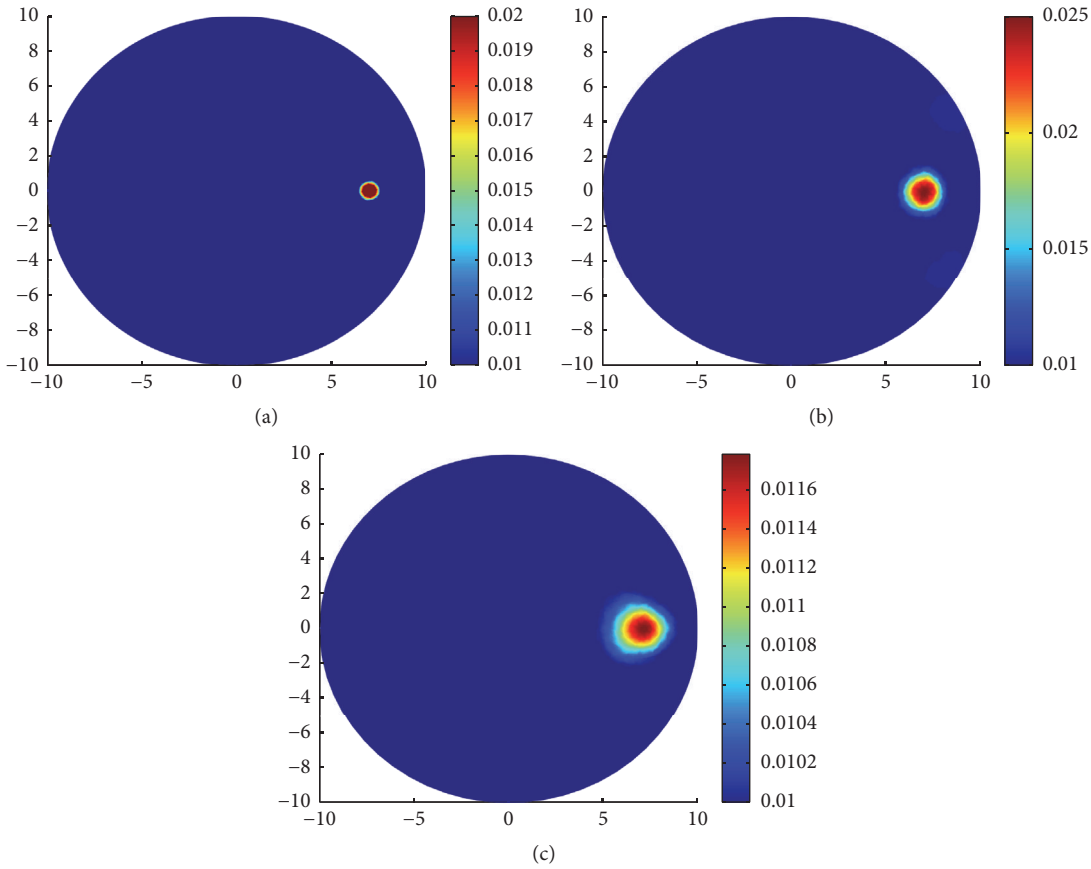


FIGURE 3: Reconstruction results with TV-L^1 regularization and TV regularization for one small inclusion. (a) True μ_a; (b) reconstruction of μ_a with TV-L^1 penalty; (c) reconstruction of μ_a with only TV penalty.

TABLE 1: Parameters corresponding to various regularization.

Cases	Small inclusion	Middle inclusion	Big inclusion
(α, β, η) for TV-L^1	$(0.0003, 0.003, 0.0003)$	$(0.003, 0.003, 0.0003)$	$(0.003, 0.0003, 0.0003)$
(α, β, η) for TV	$(3 \times 10^{-4}, 0, 0)$	$(3 \times 10^{-4}, 0, 0)$	$(3 \times 10^{-4}, 0, 0)$

for TV-L^1 regularization and Figures 3(c) and 4(c) for TV regularization. In fact, these results are reasonable and can be interpreted. The TV-L^1 regularization imposes both TV penalty and L^1 penalty on μ_a. The TV penalty tends to find edges and the L^1 penalty tends to find the sparse details of the inclusion. As can be seen from Table 1, for the big inclusion, we enhance the weight of TV penalty. When the inclusion gets smaller, we enhance the weight of L^1 penalty.

For the big inclusion, TV-L^1 regularization and TV regularization perform no big differences, but we still can see that the blurred parts in Figure 5(b) with TV-L^1 regularization are smaller than that in Figure 5(c) with TV regularization. We can see from Figures 5(b) and 5(c) that there are large area of blurs in the big inclusion case. It is reasonable in the sense that large area of absorption inclusion means more photons are absorbed in propagation process. Hence, we can alleviate this phenomenon by increasing the source-detector pairs. In other words, by increasing the measurements, one can alleviate the effect of absorption to some extent.

In the second group of experiments of this section, three types of complicated and multi-inclusions are designed for comparing the reconstruction results by TV-L^1 regularization, TV regularization, and L^1 regularization.

In the first case, we reconstruct two small circle inclusions centered at (7 mm, 0 mm) and (0 mm, 0 mm) with the same radius 0.5 mm. In both of the two inclusions, $\mu_a = 0.02\,\text{mm}^{-1}$ and $\mu_s = 20\,\text{mm}^{-1}$. We use a mesh of 1277 nodes and 2488 triangular elements for the inverse problem and a mesh of 1861 nodes and 3616 triangular elements for the forward problem. The true distributions of μ_a and reconstruction results with various regularizations can be seen in Figure 6.

In the second case, we reconstruct three small circles centered at (7 mm, 0 mm), (0 mm, −7 mm), and (0 mm, 0 mm) with the same radius, 0.5 mm. In all the three inclusions, $\mu_a = 0.02\,\text{mm}^{-1}$ and $\mu_s = 20\,\text{mm}^{-1}$. We use a mesh of 1671 nodes and 3264 triangular elements for the inverse problem and a mesh of 2141 nodes and 4176 triangulations for the forward

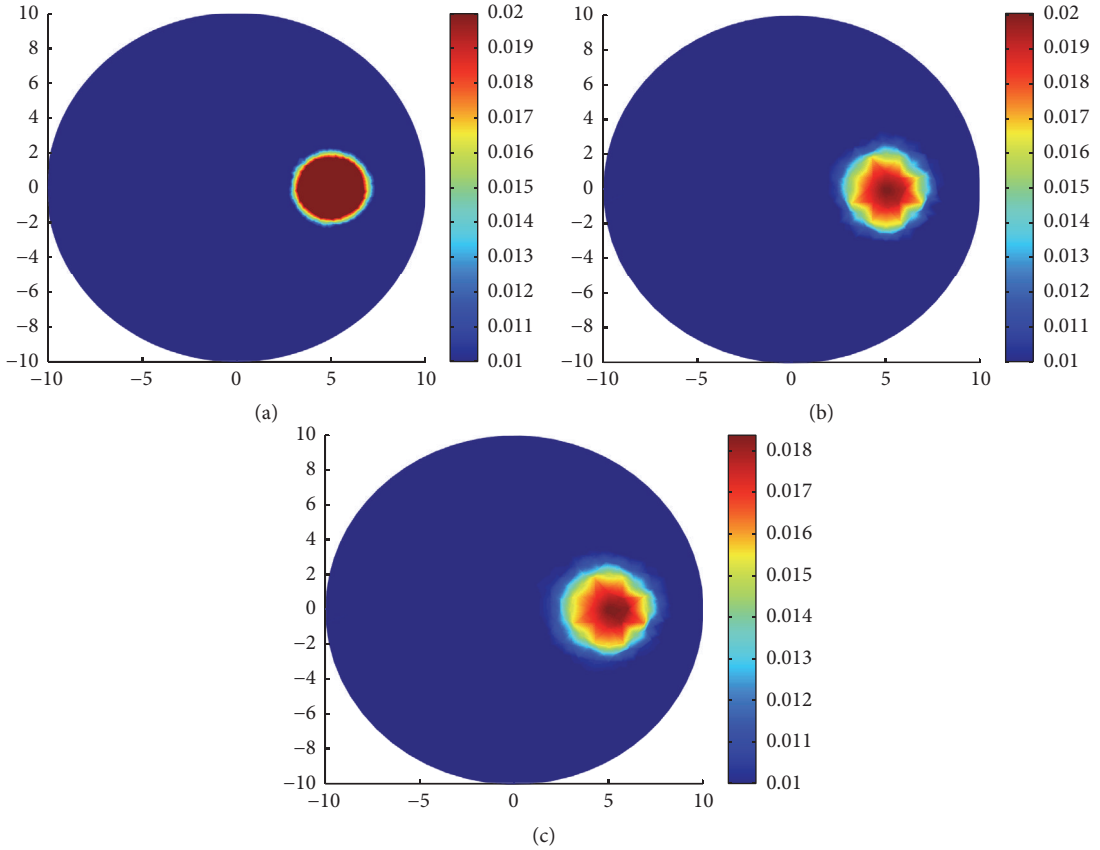


FIGURE 4: Reconstruction results with TV-L^1 regularization and TV regularization for one middle inclusion. (a) True μ_a; (b) reconstruction of μ_a with TV-L^1 penalty; (c) reconstruction of μ_a with only TV penalty.

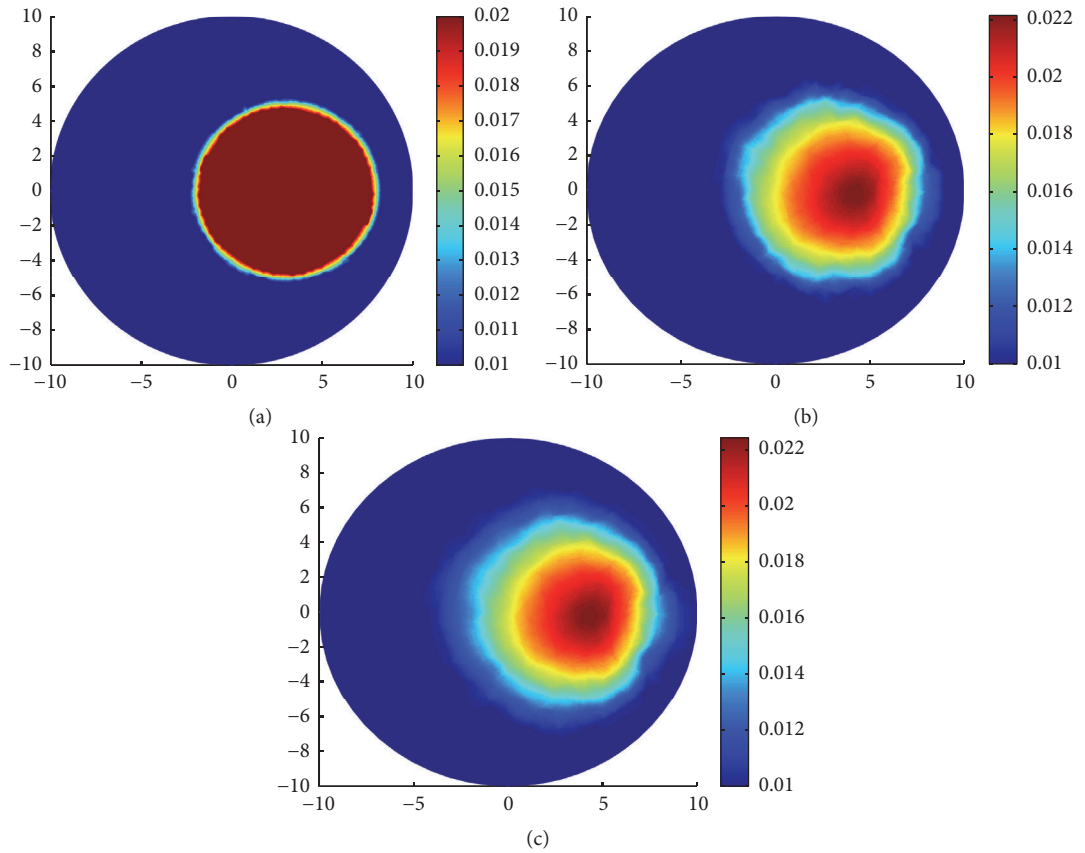


FIGURE 5: Reconstruction results with TV-L^1 regularization and TV regularization for one big inclusion. (a) True μ_a; (b) reconstruction of μ_a with TV-L^1 penalty; (c) reconstruction of μ_a with only TV penalty.

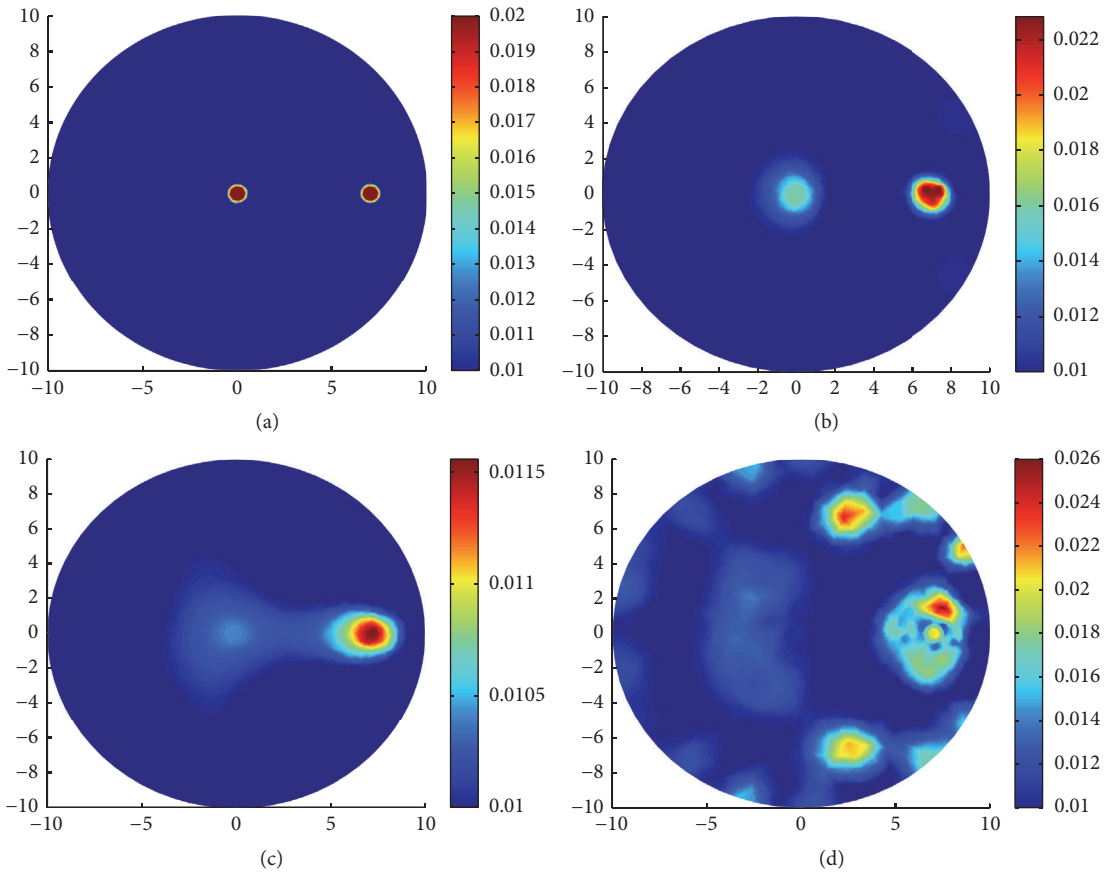


FIGURE 6: Reconstruction results with various regularization for two small inclusions. (a) True μ_a; (b) reconstruction of μ_a with TV-L^1 penalty; (c) reconstruction of μ_a with TV penalty; (d) reconstruction of μ_a with L^1 penalty.

problem. The true distributions of μ_a and reconstruction results with various regularizations can be seen in Figure 7.

In the third case, we design three inclusions. One big circle centered at (−5 mm, 2 mm) with radius 2 mm. Two small circles centered at (0 mm, −7 mm) and (7 mm, 0 mm) with the same radius 0.5 mm. In all the three cases, we take the same coefficient values $\mu_a = 0.02\,\text{mm}^{-1}$ and $\mu_s = 20\,\text{mm}^{-1}$. We use a mesh of 1323 nodes and 2568 triangulations for the inverse problem and a mesh of 1809 nodes and 3512 triangulations for the forward problem. The true distributions of μ_a and reconstruction results with various regularizations can be seen in Figure 8.

In Table 2, for all the three cases, we present the parameter settings for the iterative procedure (α, β, η), the steps of the iterations Iters, the relative residual error E_{resi}, and the relative solution error E_{μ_a} in the sense of L^2 norm corresponding to each regularization method. E_{resi} and E_{μ_a} are defined as

$$E_{\text{resi}} = \frac{\left\|F\left(\mu_a^n\right) - F\left(\mu_a^{\text{true}}\right)\right\|}{\left\|F\left(\mu_a^{\text{true}}\right)\right\|},$$
$$E_{\mu_a} = \frac{\left\|\mu_a^n - \mu_a^{\text{true}}\right\|}{\left\|\mu_a^{\text{true}}\right\|}. \tag{40}$$

We say that in the first and the second case, compared with the inclusions near the boundary, the inclusion deep in the domain is difficult to identify. In the third case, it is more difficult to reconstruct the small inclusion than the big inclusion. From the reconstruction results, we can see that the TV-L^1 regularization can preferably identify the location and the value of the inclusions. In other words, the advantages with TV-L^1 regularization are indeed apparent over TV regularization and L^1 regularization; see Figures 6(b), 7(b), and 8(b) for the reconstructions results with TV-L^1 regularization. The LM type algorithm for the TV regularization converges slowly especially when the inclusions are very small; see Figures 6(c) and 7(c). When big inclusion is included, the LM for TV regularization can well identify it, but it still cannot perfectly reconstruct the value of the small inclusions; see Figure 8(c). The split Bregman algorithm for L^1 regularization is an efficient algorithm with high convergent rate; we can see this from the number of iterations in Table 2. But, in our simulations, although the algorithm converges quickly, it loses validity when the inclusions are very small; see Figures 6(d) and 7(d). In Figure 8(d), the big inclusion is identified, but we only can see two blurred circles at where the two small inclusions are located. Moreover, we

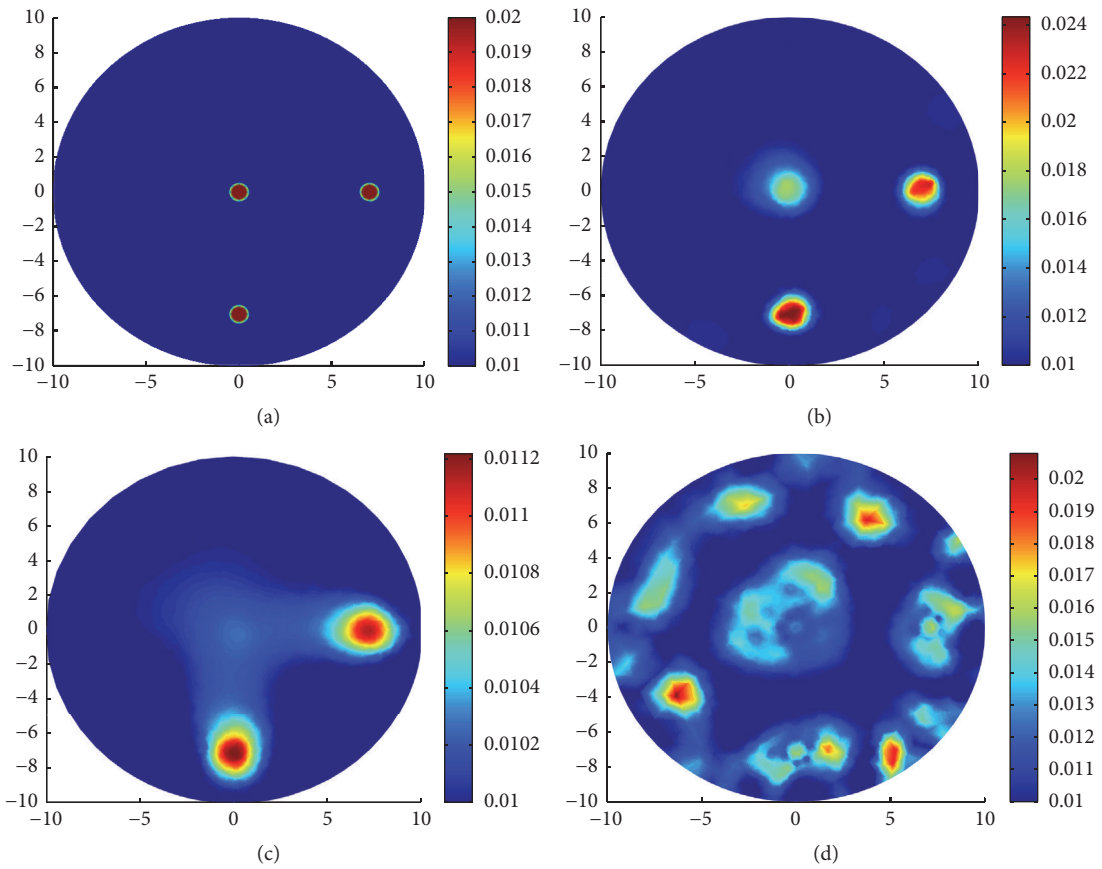


FIGURE 7: Reconstruction results with various regularization for three small inclusions. (a) True μ_a; (b) reconstruction of μ_a with TV-L^1 penalty; (c) reconstruction of μ_a with TV penalty; (d) reconstruction of μ_a with L^1 penalty.

can find that the reconstruction may be oversparsified by using the L^1 regularization.

5.3. Simulation 3: High Absorbing and Low Scattering Inclusions with TV-L^1. As the last simulation, our purpose is to show the validity of our algorithm in the situation that high absorbing inclusion and low scattering inclusion are contained in the domain X. Moreover, we investigated the reconstruction under different noise level to show the robustness of the proposed algorithm. The noises added to the exact data M_i^{true} are by the following rule: $M_i = M_i^{\text{true}}(1 + \delta * \text{random})$, where δ is the signal-to-noise ratio; random is a Gaussian random variable with zero mean and unity variation. The background values of the absorption and the scattering are $0.05\,\text{mm}^{-1}$ and $5\,\text{mm}^{-1}$, respectively. The absorbing inclusion for which we set the absorption coefficient is $\mu_a = 0.1\,\text{mm}^{-1}$ which is centered at (3.5 mm, 3.5 mm) with radius 2 mm. The scattering inclusion for which we set the scattering coefficient is $\mu_s = 10\,\text{mm}^{-1}$ which is centered at (0 mm, −5 mm) with radius 2 mm. For the discretization, we use a mesh of 463 nodes and 856 triangulations for the inverse problem and a mesh of 681 nodes and 1288 triangulations for the forward problem. See the first row of the Figure 9(a) for the true distribution of μ_a. In Figure 9(b), the reconstruction of μ_a with exact data is presented. The parameter value in this case is chosen as $\alpha = 0.0005$, $\beta = 0.0005$, and $\eta = 10^{-6}$. Figure 9(c) presents the reconstruction with noise level 0.1%, in which the parameter is taken as $\alpha = 0.0005$, $\beta = 0.0005$, and $\eta = 10^{-6}$. In Figure 9(d), the reconstruction with noise level 1% is shown, in which the parameter value is taken as $\alpha = 0.005$, $\beta = 0.0005$, and $\eta = 10^{-6}$. From the four reconstructions under four different noise levels, we can find that as the noise level increased, the quality of the reconstruction gets worse. Here, we take the noise level under 1%, since the iteration will not converge if the noise level is bigger than 1% which reflects the severe ill-posedness of the problem. We also find that in order to obtain the reconstruction results with the relative error in L^2 norm no more than 20% the noise level on the synthetic data should be less than 1%.

6. Summary

In this paper, forward and inverse problems of the radiative transfer equation are considered. An image reconstruction method based on the TV-L^1 regularization is proposed. The forward problem is discretized with the discontinuous Galerkin method on the spatial space and the finite element method on the angular space which both are implemented on the piecewise linear basis. We discretize the absorption and

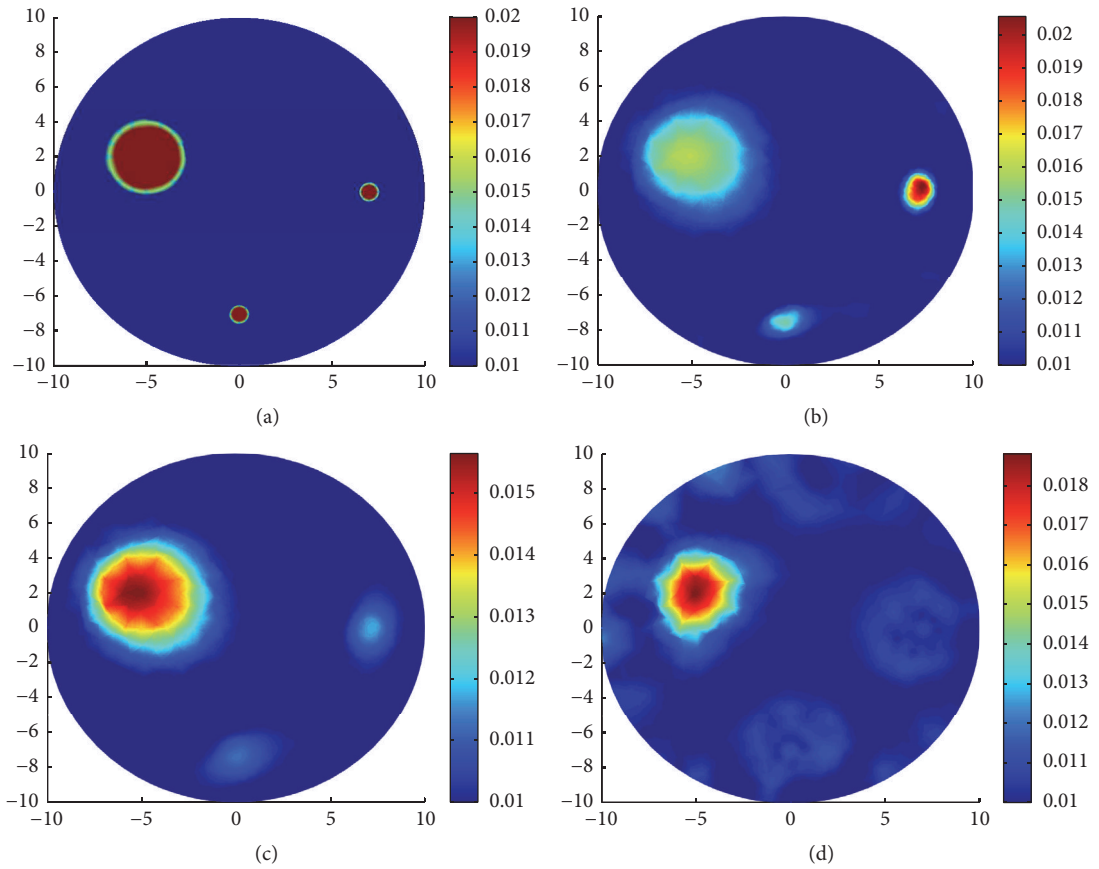


FIGURE 8: Reconstruction results with various regularization for one big and two small inclusions. (a) True μ_a; (b) reconstruction of μ_a with TV-L^1 penalty; (c) reconstruction of μ_a with TV penalty; (d) reconstruction of μ_a with L^1 penalty.

TABLE 2: Regularization parameters, error estimates, and iteration number for different cases in the second example in Section 5.2.

Cases	Parameters	TV-L^1	TV	L^1
Two small inclusions	(α, β, η)	$(10^{-4}, 10^{-3}, 10^{-5})$	$(10^{-4}, 0, 0)$	$(0, 10^{-3}, 10^{-5})$
	Iters	11	39	5
	E_{resi}	7.3×10^{-5}	6.4×10^{-5}	3.3×10^{-4}
	E_{μ_a}	13.29%	14.78%	15.76%
Three small inclusions	(α, β, η)	$(10^{-4}, 10^{-3}, 10^{-5})$	$(10^{-3}, 0, 0)$	$(0, 10^{-3}, 10^{-5})$
	Iters	8	79	6
	E_{resi}	8.6×10^{-5}	7.0×10^{-5}	6.3×10^{-4}
	E_{μ_a}	15.02%	15.80%	15.88%
One big and two small inclusions	(α, β, η)	$(10^{-3}, 10^{-3}, 10^{-5})$	$(10^{-3}, 0, 0)$	$(0, 10^{-4}, 10^{-6})$
	Iters	32	53	6
	E_{resi}	9.5×10^{-4}	6.8×10^{-4}	0.0019
	E_{μ_a}	15.75%	15.47%	15.17%

scattering coefficients on the piecewise constant basis. The minimization problem is solved by a Levenberg-Marquardt type method which is equipped with a split Bregman algorithm for the L^1 penalty. The adjoint method is used to compute the Jacobian matrix which is the discretized Fréchet derivative of the forward operator. We numerically compare the proposed reconstruction method with the other imaging reconstruction methods based on TV and L^1 regularizations. The simulation results show the validity and robustness of the proposed method.

Competing Interests

The authors declare that they have no competing interests.

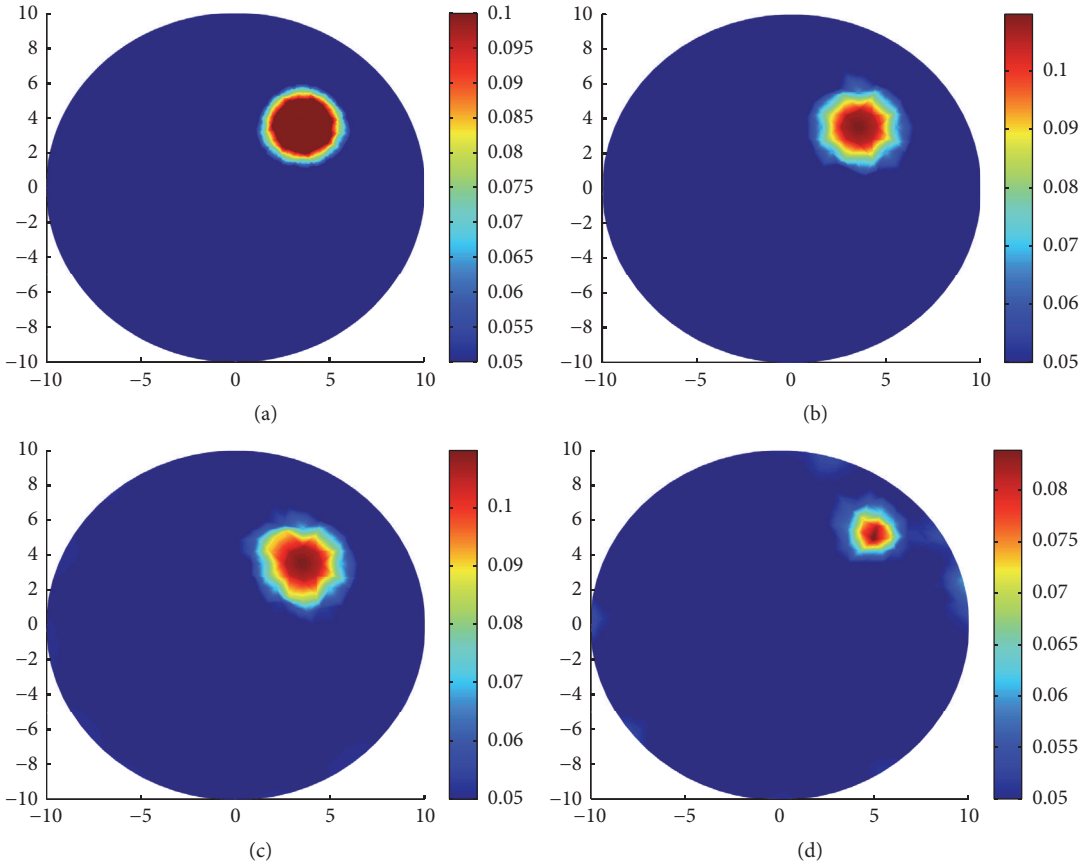


FIGURE 9: Reconstruction of μ_a with Algorithm 1 for high absorption and low scattering situation. (a) True μ_a; (b) reconstruction of μ_a with exact data; (c) reconstruction of μ_a with noise level 0.1%; (d) reconstruction of μ_a with noise level 1%.

Acknowledgments

This work is supported by the National Nature Science Foundation of China under Grants nos. 41474102 and 61307023.

References

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

[2] E. M. C. Hillman, "Optical brain imaging in vivo: techniques and applications from animal to man," *Journal of Biomedical Optics*, vol. 12, no. 5, Article ID 051402, 2007.

[3] M. Herranz and A. Ruibal, "Optical imaging in breast cancer diagnosis: the next evolution," *Journal of Oncology*, vol. 2012, Article ID 863747, 10 pages, 2012.

[4] K. Peng, L. He, B. Wang, and J. Y. Xiao, "Detection of cervical cancer based on photoacoustic imaging—the in-vitro results," *Biomedical Optics Express*, vol. 6, no. 1, pp. 135–143, 2015.

[5] H. B. Jiang, *Diffuse Optical Tomography: Principles and Applications*, CRC Press, Boca Raton, Fla, USA, 2011.

[6] S. R. Arridge and M. Schweiger, "A gradient-based optimisation scheme for optical tomography," *Optics Express*, vol. 2, no. 6, pp. 213–226, 1998.

[7] H. Dehghani, M. E. Eames, P. K. Yalavarthy et al., "Near infrared optical tomography using NIRFAST: algorithm for numerical model and image reconstruction," *Communications in Numerical Methods in Engineering with Biomedical Applications*, vol. 25, no. 6, pp. 711–732, 2009.

[8] H. B. Jiang, K. D. Paulsen, U. L. Osterberg, B. W. Pogue, and M. S. Patterson, "Optical image reconstruction using frequency-domain data: simulations and experiments," *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, vol. 13, no. 2, pp. 253–266, 1996.

[9] J. Prakash, C. B. Shaw, R. Manjappa, R. Kanhirodan, and P. K. Yalavarthy, "Sparse recovery methods hold promise for diffuse optical tomographic image reconstruction," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 20, no. 2, pp. 74–82, 2014.

[10] B. Bi, B. Han, W. Han, J. Tang, and L. Li, "Image reconstruction for diffuse optical tomography based on radiative transfer equation," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 286161, 23 pages, 2015.

[11] C. C. Leng, D. D. Yu, S. Zhang, Y. An, and Y. F. Hu, "Reconstruction method for optical tomography based on the linearized Bregman iteration with sparse regularization," *Computational and Mathematical Methods in Medicine*, vol. 2015, Article ID 304191, 8 pages, 2015.

[12] A. Chambolle and P.-L. Lions, "Image recovery via total variation minimization and related problems," *Numerische Mathematik*, vol. 76, no. 2, pp. 167–188, 1997.

[13] A. Chambolle, "An algorithm for total variation minimization and applications," *Journal of Mathematical Imaging and Vision*, vol. 20, no. 1-2, pp. 89–97, 2004.

[14] T. Chan, A. Marquina, and P. Mulet, "High-order total variation-based image restoration," *SIAM Journal on Scientific Computing*, vol. 22, no. 2, pp. 503–516, 2000.

[15] Y. Q. Dong, T. Gorner, and S. Kunis, "An algorithm for total variation regularized photoacoustic imaging," *Advances in Computational Mathematics*, vol. 41, no. 2, pp. 423–438, 2015.

[16] H. Gao, H. K. Zhao, and S. Osher, "Bregman methods in quantitative photoacoustic tomography," *Cam Report*, vol. 3, pp. 3043–3054, 2010.

[17] L. Yao and H. B. Jiang, "Enhancing finite element-based photoacoustic tomography using total variation minimization," *Applied Optics*, vol. 50, no. 25, pp. 5031–5041, 2011.

[18] J. C. Feng, C. H. Qin, K. B. Jia et al., "Total variation regularization for bioluminescence tomography with the split Bregman method," *Applied Optics*, vol. 51, no. 19, pp. 4501–4512, 2012.

[19] H. Gao and H. K. Zhao, "Multilevel bioluminescence tomography based on radiative transfer equation Part 2: total variation and l1 data fidelity," *Optics Express*, vol. 18, no. 3, pp. 2894–2912, 2010.

[20] M. Freiberger, C. Clason, and H. Scharfetter, "Total variation regularization for nonlinear fluorescence tomography with an augmented Lagrangian splitting approach," *Applied Optics*, vol. 49, no. 19, pp. 3741–3747, 2010.

[21] G. Gong, H. Zhang, and M. Yao, "Speckle noise reduction algorithm with total variation regularization in optical coherence tomography," *Optics Express*, vol. 23, no. 19, pp. 24699–24712, 2015.

[22] M. Schweiger, S. R. Arridge, and I. Nissilä, "Gauss-Newton method for image reconstruction in diffuse optical tomography," *Physics in Medicine and Biology*, vol. 50, no. 10, pp. 2365–2386, 2005.

[23] K. Ren, G. Bal, and A. H. Hielscher, "Transport- and diffusion-based optical tomography in small domains: a comparative study," *Applied Optics*, vol. 46, no. 27, pp. 6669–6679, 2007.

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

[25] J. Tang, W. Han, and B. Han, "A theoretical study for RTE-based parameter identification problems," *Inverse Problems*, vol. 29, no. 9, Article ID 095002, 18 pages, 2013.

[26] H. Gao and H. K. Zhao, "A fast-forward solver of radiative transfer equation," *Transport Theory and Statistical Physics*, vol. 38, no. 3, pp. 149–192, 2009.

[27] W. Han, J. A. Eichholz, X. L. Cheng, and G. Wang, "A theoretical framework of x-ray dark-field tomography," *SIAM Journal on Applied Mathematics*, vol. 71, no. 5, pp. 1557–1577, 2011.

[28] W. Han, J. A. Eichholz, J. Huang, and J. Lu, "RTE-based bioluminescence tomography: a theoretical study," *Inverse Problems in Science and Engineering*, vol. 19, no. 4, pp. 435–459, 2011.

[29] W. Han, J. Huang, and J. A. Eichholz, "Discrete-ordinate discontinuous Galerkin methods for solving the radiative transfer equation," *SIAM Journal on Scientific Computing*, vol. 32, no. 2, pp. 477–497, 2010.

[30] E. E. Lewis and W. F. Miller, *Computational Methods of Neutron Transport*, Wiley-Interscience, 1984.

[31] K. M. Case and P. F. Zweifel, *Linear Transport Theory*, Addison-Wesley, Reading, Mass, USA, 1967.

[32] L. C. Evans and R. F. Gariepy, *Measure Theory and Fine Properties of Functions, Revised Edition*, Chemical Rubber Company, Boca Raton, Fla, USA, 2015.

[33] T. Goldstein and S. Osher, "The split Bregman method for L1-regularized problems," *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, pp. 323–343, 2009.

[34] J. Bush, *Bregman algorithms [M.S. thesis]*, University of California, Santa Barbara, Calif, USA, 2011.

[35] D. L. Colton and R. Kress, *Inverse Acoustic and Electromagnetic Scattering Theory*, Springer, Berlin, Germany, 2nd edition, 1998.